

University of Groningen

Improved method of expression cloning in a host cell

Meima, Roelf Benhard; Wenzel, Thibaut Jose; Sagt, Cees Maria Jacobus; de Winde, Johannes Hendrik; Kuipers, Oscar Paul; van Rooijen, Rutger Jan; Dekker, Marjolein Cornelia; Vonk, brenda

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:
2006

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Meima, R. B., Wenzel, T. J., Sagt, C. M. J., de Winde, J. H., Kuipers, O. P., van Rooijen, R. J., Dekker, M. C., & Vonk, B. (2006). Improved method of expression cloning in a host cell. (Patent No. WO2006097539).

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 September 2006 (21.09.2006)

PCT

(10) International Publication Number
WO 2006/097539 A2

(51) International Patent Classification: **Not classified**

(21) International Application Number:
PCT/EP2006/060864

(22) International Filing Date: 20 March 2006 (20.03.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
05102180.6 18 March 2005 (18.03.2005) EP
05102159.0 18 March 2005 (18.03.2005) EP
05102184.8 18 March 2005 (18.03.2005) EP
05102177.2 18 March 2005 (18.03.2005) EP
05102160.8 18 March 2005 (18.03.2005) EP

(71) Applicant (for all designated States except US): **DSM IP Assets B.V.** [NL/NL]; Het Overloon 1, NL-6411 TE Heerlen (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MEIMA, Roelf, Bernhard** [NL/NL]; Hoefslag 28, NL-3471 AH Kamerik (NL). **WENZEL, Thibaut, José** [NL/NL]; Constantijn Huygenslaan 45, NL-2332 TS Leiden (NL). **SAGT, Cornelis, Maria, Jacobus** [NL/NL]; Olympus 204, NL-3534 WE Utrecht (NL). **DE WINDE, Johannes, Hendrik** [NL/NL]; Saffraanweg 51, NL-2215 WC Voorhout (NL). **KUIPERS, Oscar, Paul** [NL/NL]; Lepelaar 95, NL-9728 XG Groningen (NL). **ROOIJEN VAN, Rutger, Jan** [NL/NL]; Albertine Agneslaan 25, NL-6713 MN Ede (NL). **DEKKER, Marjolein, Conelia** [NL/NL]; Oude Raadhuislaan 17B, NL-3054 NM Rotterdam (NL). **VONK, Brenda** [NL/NL]; Selma Lagerlöfstraat 252, NL-3207 MS Spijkenisse (NL).

(74) Agents: **PALLARD, Caroline, Chantal, Patricia** et al.; DSM Intellectual Property, PP 600-0240 P.o.box 1, NL-2600 MA Delft (NL).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report
- with (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IMPROVED METHOD OF EXPRESSION CLONING IN A HOST CELL

(57) Abstract: DSM IP Assets B.V. 24403WO IMPROVED METHOD OF EXPRESSION CLONING IN A HOST CELL 5 ABSTRACT The invention relates to a promoter DNA sequence highly suited in an improved expression cloning method for isolation of DNA sequences comprising a DNA sequence encoding a protein of interest in a host cell and to the improved expression cloning method wherein use is made of this promoter. The isolated DNA sequences are useful in processes for 10 producing a protein of interest.



WO 2006/097539 A2

IMPROVED METHOD OF EXPRESSION CLONING IN A HOST CELL

5

Field of the invention

The present invention relates to a promoter highly suited for being used in a method for the identification of DNA sequences encoding proteins of interest by expression cloning.

10

Background of the invention

An increasing number of screening methods based on expression cloning are already known. Such methods have previously successfully been used for identification of prokaryotic gene products in e.g. *Bacillus* (cf. US 4,469,791, WO2005/38024) and *E.coli* (e.g. WO 95/18219 and WO 95/34662). WO2005/38024 describes a method of screening protein secreting recombinant *Bacillus* cells. However, this method is lacking high throughput, automated detection technology. Yeast has also been used as a host for expression cloning of eukaryotic genes. Strasser *et al.* (Eur. J. Biochem. (1989)**184**: 699-706) have reported the identification of a fungal α -amylase by expression cloning of fungal genomic DNA in the yeast *Saccharomyces cerevisiae*. Similarly, WO 93/11249 reports the identification of a fungal cellulase by expression cloning of fungal cDNAs in *S.cerevisiae*. Lastly, an expression cloning method in filamentous fungus has been described (WO 99/32617).

15

20

25

In all these methods, large number of transformants would have to be screened before a transformant secreting a protein with properties of interest could be isolated. There is thus a need for an expression cloning method that would optimise the chance of detecting DNA sequences encoding secreted proteins.

Brief description of the figures

30

Figure 1. Depicted is pPhtrA-gfp-amyE. This secretion reporter vector comprises the *htrA* promoter of *B.subtilis* in operative association with GFP. This vector is representative for pPhtrB-gfp-amyE.

Figure 2. Construction of *B. subtilis* reporter strains for secretion-stress. The *htr* promoters were fused to the GFP gene, finally resulting in plasmids pPhtrA-gfp-amyE and pPhtrB-gfp-

amyE. The resulting plasmids were integrated into the chromosome of *Bacillus* host cells via single crossing-over recombination in the corresponding *htr* loci, resulting in VT210A and VT210B cells. *Cmr* represents the chloramphenicol resistance gene.

5 **Figure 3.** Depicted is pUBnpr-2, comprising the neutral protease gene *npr* (Genbank accession number K 02497) from *Bacillus amyloliquefaciens*.

Figure 4. Depicted is pUBBAG, comprising the β -glucanase gene *bag* (Genbank accession number M15674) from *Bacillus amyloliquefaciens*.

10

Figure 5. Depicted is the analysis of *Bacillus subtilis* VT210A cells transformed with either plasmid pUB110 (empty control vector) or plasmid pKTH10 (encoding *Bacillus amyloliquefaciens* α -amylase gene *amyQ*) using Fluorescence Activated Cell Sorting. On the y-axis the amount of observed events is depicted, on the x-axis the relative fluorescence is depicted. In panel A, analysis of *Bacillus subtilis* VT210A cells transformed with empty vector pUB110 is depicted. In panel B, analysis of *Bacillus subtilis* VT210A cells transformed with *amyQ* vector pKTH10 is depicted. In panel C, analysis of a mixed population of *Bacillus subtilis* VT210A cells transformed with a ratio of pUB110 and pKTH10 of 1:20 is depicted. In panel D, cell sorting of a mixed population of *Bacillus subtilis* VT210A cells transformed with both pUB110 and pKTH10 is depicted; the cell sorting limit is indicated by a dashed vertical line. Cells demonstrating higher fluorescence than this limit were sorted.

25 **Figure 6.** Depicted is the analysis of a mixed population of *Bacillus subtilis* VT210B cells transformed with either plasmid pUB110 (empty control vector) or plasmid pKTH10 (encoding *Bacillus amyloliquefaciens* α -amylase gene *amyQ*). Analysis was performed by Fluorescence Activated Cell Scanning. On the y-axis the amount of observed events is depicted, on the x-axis the relative fluorescence is depicted.

30 **Figure 7.** Depicted is the pGBFINGFP-2 vector used for expression of Green Fluorescent Protein (GFP, Chalfie, M et al., Science (1994) 263(5148): 802-805) in *A. niger*. The backbone features of the depicted FIN vectors were previously described in WO 99/32617.

35 **Figure 8.** Depicted is the pGBFIN-32 vector used for expression of phytase (PHY) (phytase gene PHY identical to FytA described in WO99/32617) in *A. niger*.

Figure 9. Depicted is the pGBFIN-40 vector with the larger part of the phytase gene and the entire glucoamylase promoter removed, this vector is used as empty control vector for construction of *A. niger* WT-vector strain.

5 **Figure 10.** Depicted is a Northern Blot analysis performed to confirm the identification of secretion-induced promoters.

Figure 11. Depicted is a vector used for gene expression in *A. niger*. This vector is used as backbone in the construction of the secretion-induced *A.niger* reporter constructs.

10

Figure 12. Depicted is the pGBFINGFPBLE-1 cloning vector. This secretion-inducible reporter construct comprises the GFP-BLE reporter fusion gene in operative association with secretion-inducible promoter 1 of the present invention. The construct is representative for pGBFINGFPBLE-2, pGBFINGFPBLE-3, pGBFINGFPBLE-4, and pGBFINGFPBLE-5.

15

Figure 13. Depicted is pGBTOPSEL-1, this vector is used as the backbone for construction of the final secretion-inducible reporter plasmids. Backbone features of the depicted vector were previously described in WO 99/32617.

20

Figure 14. Depicted is the final secretion reporter construct pGBTOPGFPBLE-1. This secretion-induced reporter vector comprises the GFP-BLE reporter construct in operative association with secretion-inducible promoter 1 of the present invention. The vector is representative for pGBTOPGFPBLE-2, pGBTOPGFPBLE-3, pGBTOPGFPBLE-4, and pGBTOPGFPBLE-5.

25

Figure 15. Depicted is the difference in phleomycin resistance of *A. niger* co-transformants containing a reporter construct with secretion-inducible promoter P4 in combination with a library construct encoding either an intracellular protein (represented by black dots) or an extracellular protein (represented by black squares).

30

Figure 16. Depicted is the difference in phleomycin resistance of *A. niger* co-transformants containing a reporter construct with secretion-inducible promoter P5 in combination with a library construct encoding either an intracellular protein (represented by black dots) or an extracellular protein (represented by black squares).

35

Figure 17. Depicted is an E-PAGE™ gel, wherein each lane represents a co-transformed *A.niger* clone comprising a reporter construct with secretion-inducible promoter P1 in combination with a library construct encoding an extracellular protein. M represents the E-PAGE SeeBlue® prestained standard (Invitrogen, U.K.).

Figure 18. Depicted is an E-PAGE™ gel, wherein each lane represents a co-transformed *A.niger* clone comprising a reporter construct with secretion-inducible promoter P3 in combination with a library construct encoding an extracellular protein. M represents the E-PAGE SeeBlue® prestained standard (Invitrogen, U.K.).

Description of the invention

Promoter DNA sequence

According to a first aspect of the invention, there is provided a promoter DNA sequence such as:

- (a) a DNA sequence as presented in the following list: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24,
- (b) a DNA sequence capable of hybridizing with a DNA sequence of (a),
- (c) a DNA sequence being at least 50% homologous to a DNA sequence of (a),
- (d) a variant of any of the DNA sequences of (a) to (c), or
- (e) a subsequence of any of the DNA sequences of (a) to (d).

According to a preferred embodiment, the promoter sequence of the invention is one promoter DNA sequence such as:

- (a) one DNA sequence as presented in the following list: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24,
- (b) one DNA sequence capable of hybridizing with a DNA sequence of (a),
- (c) one DNA sequence being at least 50% homologous to a DNA sequence of (a),
- (d) one variant of any of the DNA sequences of (a) to (c), or
- (e) one subsequence of any of the DNA sequences of (a) to (d).

In the context of this application, a promoter DNA sequence is a DNA sequence, which is capable of controlling the expression of a coding sequence, when this promoter DNA sequence is in operative association with this coding sequence. The term "in operative

association" is defined herein as a configuration in which a promoter DNA sequence is appropriately placed at a position relative to a coding sequence such that the promoter DNA sequence directs the production of a polypeptide encoded by the coding sequence.

The term "coding sequence" is defined herein as a nucleic acid sequence that is transcribed into mRNA, which is translated into a polypeptide when placed under the control of the appropriate control sequences. The boundaries of the coding sequence are generally determined by the ATG start codon, which is normally the start of the open reading frame at the 5' end of the mRNA and a transcription terminator sequence located just downstream of the open reading frame at the 3' end of the mRNA. A coding sequence can include, but is not limited to, genomic DNA, cDNA, semisynthetic, synthetic, and recombinant nucleic acid sequences. Preferably, a promoter DNA sequence is defined by being the DNA sequence located upstream of a coding sequence associated thereto and by being capable of controlling the expression of this coding sequence.

More specifically, the term "promoter" is defined herein as a DNA sequence that binds the RNA polymerase and directs the polymerase to the correct downstream transcriptional start site of a coding sequence encoding a polypeptide to initiate transcription. RNA polymerase effectively catalyzes the assembly of messenger RNA complementary to the appropriate DNA strand of the coding region. The term "promoter" will also be understood to include the 5' non-coding region (between promoter and translation start) for translation after transcription into mRNA, cis-acting transcription control elements such as enhancers, and other nucleotide sequences capable of interacting with transcription factors.

In a preferred embodiment, the promoter DNA sequence of the invention is a DNA sequence derived from: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24.

According to another preferred embodiment, the promoter DNA sequence of the invention is a DNA sequence capable of hybridizing with a DNA sequence as presented in the following list: SEQ ID NO 1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 and which still retains promoter activity.

Promoter activity is preferably determined by measuring the concentration of the protein(s) produced as a result of the expression of a coding sequence(s), which is (are) in operative association with the promoter. Alternatively the promoter activity is determined by measuring the enzymatic activity of the protein(s) coded by the coding sequence(s), which is (are) in operative association with the promoter. According to a preferred embodiment, the promoter activity (and its strength) is determined by measuring the expression of the

coding sequence of the *lacZ* reporter gene (In Luo, Gene 163 (1995) 127-131 or Perkins and Youngman (1986) *Proc Natl Acad Sci USA* **83**: 140; or Vagner *et al.* (1998) *Microbiology* **144**: 3097). According to another preferred embodiment, the promoter activity is determined by using the green fluorescent protein as coding sequence (In *Microbiology*. 1999 Mar;145 (Pt 3):729-34. Santerre Henriksen AL, Even S, Muller C, Punt PJ, van den Hondel CA, Nielsen J.Study). In bacterial cells, GFP and other fluorescent proteins have been used extensively for expression and (dynamic) protein localization studies (reviewed in Southward and Surette (2002) *Mol. Microbiol.* **45**: 1191). Additionally, promoter activity can be determined by measuring the mRNA levels of the transcript generated under control of the promoter. The mRNA levels can, for example, be measured by Northern blot, GeneChips™(Affymetrix) or spotted array technology or quantitative (real time) PCR (J. Sambrook, E. F. Fritsch, and T. Maniatis, 2001, *Molecular Cloning, A Laboratory Manual*, 3d edition, Cold Spring Harbor, N.Y.). Most preferably, promoter activity is determined by mRNA analysis using Northern blot.

Preferably, (isolated) promoter DNA sequences of the invention hybridize under very low stringency conditions, more preferably low stringency conditions, more preferably medium stringency conditions, more preferably medium-high stringency conditions, even more preferably high stringency conditions, and most preferably very high stringency conditions with a nucleic acid probe which hybridizes under the same conditions with:

- (i) SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24., or
- (ii) a subsequence of (i), or
- (iii) a complementary strand of (i), (ii),

The term complementary strand is known to the person skilled in the art and is described in (J. Sambrook, E. F. Fritsch, and T. Maniatus, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, N.Y.). A subsequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 may be ranged between 20 nucleotides and the respective whole sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 upstream the coding region (the corresponding coding regions of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 are respectively given in SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28 and SEQ ID NO:29).

The nucleic acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 or a subsequence thereof (as defined in the previous paragraph) may be used to design a nucleic acid probe to identify

and clone DNA promoters from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic or cDNA of the genus or species of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 15, preferably at least 25, and more preferably at least 35 nucleotides in length. Additionally, such probes can be used to amplify DNA promoters through PCR. Longer probes can also be used. DNA, RNA and Peptide Nucleic Acid (PNA) probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with 32 P, 33 P, 3 H, 35 S, biotin, or avidin or a fluorescent marker). Such probes are encompassed by the present invention.

Thus, a genomic DNA or cDNA library prepared from such other organisms may be screened for DNA, which hybridizes with the probes described above and which encodes a polypeptide. Genomic or other DNA from such other organisms may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA sequence which is homologous with SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 or a subsequence thereof, the carrier material may be used in a Southern blot. For purposes of the present invention, hybridization indicates that the DNA sequence hybridizes under very low to very high stringency conditions to a labeled nucleic acid probe corresponding to the DNA sequence shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24, to the complementary strands of said DNA sequences, or to subsequences of said DNA sequences. Molecules to which the nucleic acid probe hybridizes under these conditions are detected using for example a X-ray film. Other hybridisation techniques also can be used, such as techniques using fluorescence for detection and glass slides and/or DNA microarrays as support. An example of DNA microarray hybridisation detection is given in FEMS Yeast Res. 2003 Dec;4(3):259-69 (Daran-Lapujade P, Daran JM, Kotter P, Petit T, Piper MD, Pronk JT. "Comparative genotyping of the *Saccharomyces cerevisiae* laboratory strains S288C and CEN.PK113-7D using oligonucleotide microarrays". Additionally, the use of PNA microarrays for hybridization is described in Nucleic Acids Res. 2003 Oct 1;31(19):e119 (Brandt O, Feldner J, Stephan A, Schroder M, Schnolzer M, Arlinghaus HF, Hoheisel JD, Jacob A. PNA microarrays for hybridisation of unlabelled DNA samples.)

In a preferred embodiment, the nucleic acid or DNA probe is the DNA sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24. In another embodiment, the nucleic acid probe is the DNA sequence having nucleotides 100 to 150 of SEQ ID NO:1 or the DNA sequence having nucleotides 170 to 220 of SEQ ID NO:2.

In another preferred embodiment, the nucleic acid probe is the DNA sequence having :

- nucleotides 996 to 2797 of SEQ ID NO: 20, nucleotides 996 to 3561 of SEQ ID NO: 21, nucleotides 997 to 3536 of SEQ ID NO: 22, nucleotides 1034 to 3175 of SEQ ID NO: 23, or nucleotides 1080 to 4054 of SEQ ID NO: 24, more preferably:
- nucleotides 1496 to 2497 of SEQ ID NO: 20, nucleotides 1496 to 3261 of SEQ ID NO: 21, nucleotides 1497 to 3236 of SEQ ID NO: 22, nucleotides 1534 to 2875 of SEQ ID NO: 23, or nucleotides 1580 to 3754 of SEQ ID NO: 24, even more preferably:
- nucleotides 1896 to 2197 of SEQ ID NO: 20, nucleotides 1896 to 2861 of SEQ ID NO: 21, nucleotides 1897 to 2936 of SEQ ID NO: 22, nucleotides 1934 to 2575 of SEQ ID NO: 23, or nucleotides 1980 to 3454 of SEQ ID NO: 24, and most preferably:
- nucleotides 1996 to 2100 of SEQ ID NO: 20, nucleotides 1996 to 2300 of SEQ ID NO: 21, nucleotides 1997 to 2400 of SEQ ID NO: 22, nucleotides 2034 to 2100 of SEQ ID NO: 23, or nucleotides 2080 to 2800 of SEQ ID NO: 24.

According to another preferred embodiment, the probe is part of the DNA sequence of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 located upstream of the coding sequence. The corresponding coding regions of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 are respectively given in SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28 and SEQ ID NO:29).

For long probes of at least 100 nucleotides in length, very low to very high stringency conditions are defined as prehybridization and hybridization at 42 degrees Celsius in 5 times SSPE, 0.3% SDS, 200 microgram/ml sheared and denatured salmon sperm DNA, and either 25% formamide for very low and low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures. For long probes of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 2 times SSC, 0.2% SDS preferably at least at 45 DEG C. (very low stringency), more preferably at least at 50 degrees Celsius (low stringency), more preferably at least at 55 degrees Celsius (medium stringency), more preferably at least at 60 degrees Celsius (medium-high stringency), even more preferably at least at 65 degrees Celsius (high stringency), and most preferably at least at 70 degrees Celsius (very high stringency).

For short probes which are about 15 nucleotides to about 70 nucleotides in length, stringency conditions are defined as prehybridization, hybridization, and washing post-hybridization at 5 degrees Celsius to 10 degrees Celsius below the calculated T_m using the calculation according to Bolton and McCarthy (1962, Proceedings of the National Academy of Sciences USA 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1.times.Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2 mg of yeast RNA per ml following standard Southern blotting procedures. For short probes, which are about 15 nucleotides to about 70 nucleotides in length, the carrier material is washed once in 6 times SCC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6 times SSC at 5 degrees Celsius to 10 degrees Celsius below the calculated T_m.

According to another preferred embodiment, the promoter DNA sequence of the invention derived from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 is first used to clone the native gene, coding sequence or part of it, which is operatively associated with it. This can be done starting with either SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 or a subsequence thereof as earlier defined and using this sequence as a probe. The probe is hybridised to a cDNA or a genomic library of a given host either *Bacillus*, *Aspergillus niger* or any other host as defined in this application. Once the native gene or part of it has been cloned, it can be subsequently used itself as a probe to clone homologous genes thereof derived from other host by hybridisation experiments as described herein. In this context, a homologous gene means a gene, which is at least 50% homologous to the native gene. Preferably, the homologous gene is at least 55% homologous, more preferably at least 60%, more preferably at least 65%, more preferably at least 70%, even more preferably at least 75% preferably about 80%, more preferably about 90%, even more preferably about 95%, and most preferably about 97% homologous to the native gene.

The sequence upstream the coding sequence of the homologous gene is a promoter encompassed by the present invention. Alternatively, the sequence of the native gene, coding sequence or part of it, which is operatively associated with a promoter of the invention can be identified by using SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28 or SEQ ID NO:29 or a subsequence thereof as earlier defined to search genomic databases using for example an alignment or BLAST algorithm as described herein. This identified sequence subsequently can be used to identify orthologues or homologous genes in any other fungal host as defined in this application. The sequence upstream the coding

sequence of the identified orthologue or homologous gene is a promoter encompassed by the present invention.

According to another preferred embodiment, the promoter DNA sequence of the invention is a(n) (isolated) DNA sequence derived from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 and which still has promoter activity as defined earlier.

According to another preferred embodiment, the promoter DNA sequence of the invention is a(n) (isolated) DNA sequence derived from either SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24, respectively, which is at least 50% homologous to at least part of the respective corresponding DNA sequences: SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19 situated upstream of the respective coding regions (SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28 or SEQ ID NO:29). Preferably, the derived DNA sequence is at least 55% homologous, more preferably at least 60%, more preferably at least 65%, more preferably at least 70%, even more preferably at least 75% preferably about 80%, more preferably about 90%, even more preferably about 95%, and most preferably about 97% homologous to at least part of SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19 situated upstream of the respective corresponding coding region.

For purposes of the present invention, the degree of homology between two nucleic acid sequences is preferably determined by the Wilbur-Lipman method (Wilbur and Lipman, 1983, Proceedings of the National Academy of Science USA 80: 726-730) using the LASERGENE.TM. MEGALIGN.TM. software (DNASTAR, Inc., Madison, Wis.) with an identity table and the following multiple alignment parameters: Gap penalty of 10 and gap length penalty of 10. Pairwise alignment parameters were Ktuple=3, gap penalty=3, and windows=20.

According to another preferred embodiment, the promoter DNA sequence of the invention is derived from a(n) (isolated) DNA sequence, which is a variant of either SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19 or part of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19 situated upstream of the respective corresponding coding region. The term "variant promoter" is defined herein as a promoter having a nucleotide sequence comprising a substitution, deletion, and/or insertion of one or more nucleotides of a parent promoter, wherein the variant promoter has more or less promoter activity than the corresponding parent promoter. The term "variant promoter" will also encompass natural variants and in vitro generated variants obtained using methods well known in the art such as classical mutagenesis, site-directed mutagenesis, and DNA

shuffling. A variant promoter may have one or more mutations. Each mutation is an independent substitution, deletion, and/or insertion of a nucleotide. The introduction of a substitution, deletion, and/or insertion of a nucleotide into the promoter may be accomplished using any of the methods known in the art such as classical mutagenesis, site-directed mutagenesis, or DNA shuffling. Particularly useful is a procedure, which utilizes a supercoiled, double stranded DNA vector with an insert of interest and two synthetic primers containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, extend during temperature cycling by means of Pfu DNA polymerase. On incorporation of the primers, a mutated plasmid containing staggered nicks is generated. Following temperature cycling, the product is treated with DpnI, which is specific for methylated and hemimethylated DNA to digest the parental DNA template and to select for mutation-containing synthesized DNA. Other procedures known in the art may also be used. Example of other procedures are the QuickChange™ site-directed mutagenesis kit (Stratagene Cloning Systems, La Jolla, CA), the 'The Altered Sites® II in vitro Mutagenesis Systems' (Promega Corporation) or by overlap extension using PCR as described in Gene. 1989 Apr 15;77(1):51-9. (Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR "Site-directed mutagenesis by overlap extension using the polymerase chain reaction") or using PCR as described in *Molecular Biology: Current Innovations and Future Trends*. (Eds. A.M. Griffin and H.G.Griffin. ISBN 1-898486-01-8 1995 Horizon Scientific Press, PO Box 1, Wymondham, Norfolk, U.K.). According to a preferred embodiment, the variant promoter is a promoter, which has at least one modified regulatory site as compared to the promoter sequence first derived from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19. Such regulatory site can have been removed in its entirety or specifically mutated as explained above. The regulation of such promoter variant is thus modified so that for example it is no longer induced by glucose. Examples of such promoter variants and techniques on how to obtain them are described in EP 673 429B or in WO 94/04673.

The promoter variant can be an allelic variant. An allelic variant denotes any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. The variant promoter may be obtained by the following method comprising steps (a) and (b): (a) hybridizing a DNA sequence under very low, low, medium, medium-high, high, or very high stringency conditions with:

- (i) SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19, or
- (ii) a subsequence of (i), or

(iii) a complementary strand of (i), (ii), and
(b) isolating the variant promoter from the DNA sequence. Stringency and wash conditions are defined herein.

The variant promoter can be a promoter, whose sequence may be provided with
5 linkers for the purpose of introducing specific restriction sites facilitating ligation of the promoter sequence with the coding region of the nucleic acid sequence encoding a polypeptide.

In another preferred embodiment, the promoter DNA sequence is derived from a
subsequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID
10 NO:17, SEQ ID NO:18 or SEQ ID NO:19. A subsequence is preferably defined in this context as being part of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19 located upstream of the respective corresponding coding sequence as defined previously.

The subsequence preferably contains at least about 50 nucleotides, or preferably at
15 least about 100 nucleotides, more preferably at least about 200 nucleotides, even more preferably at least about 230 nucleotides, even more preferably at least about 250 nucleotides, and most preferably at least about 290 nucleotides.

According to another preferred embodiment, the subsequence differs from the one
defined in former sentence by the fact that one or more nucleotides from the 5' and/or 3' end
20 have been deleted, said DNA sequence still having promoter activity as defined earlier.

In another preferred embodiment, the promoter subsequence is a 'trimmed'
subsequence from translation start and/or from transcription start. An example of trimming
a promoter and functionally analysing it is described in Gene. 1994 Aug 5;145(2):179-87: the
effect of multiple copies of the upstream region on expression of the *Aspergillus niger*
25 glucoamylase-encoding gene. Verdoes JC, Punt PJ, Stouthamer AH, van den Hondel CA).

The sequence information as provided herein should not be so narrowly construed
as to require inclusion of erroneously identified bases. The specific sequences disclosed
herein can be readily used and be eventually subjected to further sequence analyses
thereby identifying sequencing errors to isolate the original DNA sequence from a
30 filamentous fungus, in particular *Aspergillus niger*.

Unless otherwise indicated, all nucleotide sequences determined by sequencing a
DNA molecule herein were determined using an automated DNA sequencer. Therefore, as
is known in the art for any DNA sequence determined by this automated approach, any
nucleotide sequence determined herein may contain some errors. Nucleotide sequences
35 determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the

sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art.

The person skilled in the art is capable of identifying such erroneously identified bases and knows how to correct for such errors.

5 Functional nucleic acid equivalents may typically contain mutations that do not alter the biological function of the promoter it contains. The term "functional equivalents" also encompasses orthologues of the *A. niger* DNA sequences. Orthologues of the *A. niger* DNA sequences are DNA sequences that can be isolated from other strains or species and possess a similar or identical biological activity.

10 Homologous (similar or identical) sequences can also be determined by using a "sequence comparison algorithm". Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2: 482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48: 443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Nat'l Acad. Sci. USA 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by visual inspection. An example of an algorithm that is suitable for determining sequence similarity is the BLAST algorithm, which is described in Altschul, et al., J. Mol. Biol. 215: 403-410 (1990). Alternatively other programs can be used for sequence alignment as earlier described herein. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. These initial neighborhood word hits act as starting points to find longer HSPs containing them. The word hits are expanded in both directions along each of the two sequences being compared for as far as the cumulative alignment score can be increased. Extension of the word hits is stopped when: the cumulative alignment score falls off by the quantity X from a maximum achieved value; 20 the cumulative score goes to zero or below ; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a wordlength (W) of 11, the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89: 10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands. 30 The BLAST algorithm then performs a statistical analysis of the similarity between two sequences (see, e. g., Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90: 5873-5787 (1993)).

One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P (N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, an amino acid sequence is considered similar to a protein such as a protease if the smallest sum probability in a comparison of the test amino acid sequence to a protein such as a protease amino acid sequence is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

Preferably the similarity of the variant promoter is at least 40% homology to one of the DNA sequences having SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19. More preferably the similarity is at least 50%, more preferably, at least 60%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90% and more preferably at least 95% or at least 98% homology to the DNA sequence having SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19.

In addition to naturally occurring allelic variants of the promoter sequence, the skilled person will recognise that changes can be introduced by mutation into the promoter sequence derived from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19, without substantially altering its promoter function.

The promoter sequences of the present invention may be obtained from microorganisms of any genus. For purposes of the present invention, the term "obtained from" as used herein in connection with a given source shall mean that the polypeptide is produced by the source or by a cell in which a gene from the source has been inserted.

Preferably, the microorganism is a prokaryote. Preferred prokaryotes are *Bacillus* and *E. coli*. Preferred *Bacilli* are *Bacillus subtilis*, *Bacillus amyloliquefaciens*, *Bacillus licheniformis*, *Bacillus alcalophilus*, *Bacillus clausii*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus firmus*, *Bacillus pumilis*, *Bacillus stearothermophilus*, *Bacillus megaterium*, *Bacillus lentus*, or *Bacillus thuringiensis*.

Alternatively, the promoter sequence is obtained from an eukaryote, preferably a fungal source, and more preferably from a yeast strain such as a *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* strain; or more preferably from a filamentous fungal strain such as an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, or *Trichoderma* strain.

In a preferred embodiment, the promoter sequences are obtained from a *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, *Saccharomyces bayanus*, *Saccharomyces bayanus*, var. *uvarum* or *Saccharomyces pastorianus* strain.

In another preferred embodiment, the promoter sequences are obtained from an *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *A. nidulans*, *A. niger*, preferably *A. niger* CBS 513.88, *A. sojae*, *Aspergillus oryzae* (*A. oryzae*), *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* strain.

In another preferred embodiment, the promoter sequences are obtained from a *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum* strain.

It will be understood that for the aforementioned species, the invention encompasses the perfect and imperfect states, and other taxonomic equivalents, e.g., anamorphs, regardless of the species name by which they are known. Those skilled in the art will readily recognize the identity of appropriate equivalents. Strains of these species are readily accessible to the public in a number of culture collections, such as the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

Furthermore, such nucleic acid sequences may be identified and obtained from other sources including microorganisms isolated from nature (e.g. soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms from natural habitats are well known in the art. The nucleic acid sequence may then be derived by similarly screening a genomic DNA library of another microorganism. Once a nucleic acid sequence encoding a promoter has been detected with the probe(s), the sequence may be isolated or cloned by utilizing techniques which are known to those of ordinary skill in the art (see, e.g., Sambrook et al., 1989, supra).

In the present invention, the promoter DNA sequence may also be a hybrid promoter comprising a portion of one or more promoters of the present invention; a portion of a promoter of the present invention and a portion of another known promoter, e.g., a leader

sequence of one promoter and the transcription start site from the other promoter; or a portion of one or more promoters of the present invention and a portion of one or more other promoters. The other promoter may be any promoter sequence, which shows transcriptional activity in the host cell of choice including a variant, truncated, and hybrid promoter, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell. The other promoter sequence may be native or foreign to the nucleic acid sequence encoding the polypeptide and native or foreign to the cell.

As a preferred embodiment, important regulatory subsequences of the promoter identified can be fused to other 'basic' promoters to enhance their promoter activity (as for example described in Mol Microbiol. 1994 May;12(3):479-90. Regulation of the xylanase-encoding xlnA gene of *Aspergillus tubigenensis*. de Graaff LH, van den Broeck HC, van Ooijen AJ, Visser J.).

Other examples of other promoters useful in the construction of hybrid promoters with the promoters of the present invention include the promoters obtained from the genes for *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral alpha-amylase, *A. niger* acid stable alpha-amylase, *A. niger* or *Aspergillus awamori* glucoamylase (glaA), *A. niger* gpdA, *A. niger* glucose oxidase goxC, *Rhizomucor miehei* lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase, *A. nidulans* acetamidase, and *Fusarium oxysporum* trypsin-like protease (WO 96/00787), as well as the NA2-tpi promoter (a hybrid of the promoters from the genes for *A. niger* neutral alpha-amylase and *A. oryzae* triose phosphate isomerase), *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, and mutant, truncated, and hybrid promoters thereof. Other useful promoters for yeast host cells are described by Romanos et al., 1992, Yeast 8: 423-488. Other useful promoters for bacterial cells are the promoter of the following genes *B. subtilis* alkaline protease (*apr*), *B. subtilis* neutral protease (*npr*), *B. amyloliquefaciens* α -amylase (*amyQ*), *B. amyloliquefaciens* alkaline protease (*apr*), and *B. amyloliquefaciens* neutral protease (*npr*).

In the present invention, the promoter DNA sequence may also be a "tandem promoter". A "tandem promoter" is defined herein as two or more promoter sequences each of which is in operative association with a coding sequence and mediates the transcription of the coding sequence into mRNA.

The tandem promoter comprises two or more promoters of the present invention or alternatively one or more promoters of the present invention and one or more other known

promoters, such as those exemplified above useful for the construction of hybrid promoters. The two or more promoter sequences of the tandem promoter may simultaneously promote the transcription of the nucleic acid sequence. Alternatively, one or more of the promoter sequences of the tandem promoter may promote the transcription of the nucleic acid sequence at different stages of growth of the cell or, in the case of fungal hosts, morphological different parts of the mycelia, or different cellular compartments during spore development in sporulating bacteria such as bacilli.

In the present invention, the promoter may be foreign to the coding sequence and/or to the host cell. A variant, hybrid, or tandem promoter of the present invention will be understood to be foreign to a DNA sequence encoding a polypeptide even if the wild-type promoter is native to the coding sequence or to the host cell.

A variant, hybrid, or tandem promoter of the present invention has at least about 20%, preferably at least about 40%, more preferably at least about 60%, more preferably at least about 80%, more preferably at least about 90%, more preferably at least about 100%, even more preferably at least about 200%, most preferably at least about 300%, and even most preferably at least about 400% of the promoter activity of any promoter derived from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24 and defined earlier. The promoter activity is preferably determined as defined earlier. Most preferably, promoter activity is determined by mRNA analysis using Northern blot.

DNA construct

According to a second aspect, the invention provides a DNA construct comprising a promoter DNA sequence of the invention as defined in the former section operatively associated with a reporter gene conferring a selectable trait. The reporter gene can be any gene conferring a selectable trait to a suitable host. Preferably, the reporter gene is a selection marker gene.

"DNA construct" is defined herein as a nucleic acid molecule, either single or double-stranded, which is isolated from a naturally occurring gene or which has been modified to contain segments of nucleic acid combined and juxtaposed in a manner that would not otherwise exist in nature.

The term DNA construct is synonymous with the term expression cassette when the DNA construct contains a coding sequence or reporter gene and all the control sequences required for expression of the coding sequence or reporter gene. In the context of this application the term "DNA construct" is interchangeably used with the term "isolated DNA construct" wherein both terms are construed to be synonymous.

The DNA construct may comprise a selection marker gene, which permits easy selection of transformed host cells. In case the DNA construct comprises two selection marker genes, the skilled person will understand that these two selection marker genes are not identical: the first one associated with the promoter of the invention is used in the screening method of the invention and the second one present on the DNA construct allows a standard selection of transformed host cells.

The reporter gene is in operative association with the promoter DNA sequence of the invention such that the reporter gene can be expressed under the control of the promoter DNA sequence in a given host cell. The polypeptide encoded by the reporter gene may be native or heterologous to the host cell and/or to the promoter DNA sequence.

In this context "a" means "at least one". Therefore, the DNA construct comprises at least one reporter gene. The host may be co-transformed with at least two DNA constructs, one comprising the selection marker and another one comprising the promoter of the invention in association with a reporter gene.

Alternatively, and according to a more preferred embodiment, the reporter gene comprised in the DNA construct is a hybrid reporter gene comprising at least part of a first reporter gene with at least part of a second reporter gene, wherein the coding sequences of the first and second reporter genes are in-frame coupled to each other and wherein said hybrid reporter gene is operably linked to the promoter of the invention. The first and/or second reporter gene may be any gene conferring a selectable trait to a suitable host. A preferred hybrid reporter gene for fungi comprises the GFP gene and the phleomycin resistance gene (BLE) of *Streptoalloteichus hindustanis* (Drocourt *et al.*, NAR, 1990).

A selection marker or selectable marker is a gene, the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like. Suitable selectable markers for prokaryotic host cells include, but are not limited to, kanamycin, neomycin, erythromycin and other MLS-type markers (macrolide- lincosamide-streptogramin B), chloramphenicol, ampicillin, tetracyclin, streptomycin and spectinomycin. Suitable markers for yeast host cells are ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. Selectable markers for use in a filamentous fungal host cell include, but are not limited to, amdS (acetamidase), argB (ornithine carbamoyltransferase), bar (phosphinothricin acetyltransferase), hygB (hygromycin phosphotransferase), niaD (nitrate reductase), pyrG (orotidine-5'-phosphate decarboxylase), sC (sulfate adenylyltransferase), trpC (anthranilate synthase), as well as equivalents thereof. Marker conferring resistance against e.g. phleomycin, hygromycin B or G418 can also be used. Preferred for use in an *Aspergillus* cell are the amdS and pyrG genes of *A. nidulans* or *A. oryzae* and the bar gene of *Streptomyces hygroscopicus*. The amdS marker gene is preferably used applying the

technique described in EP 635 574B or WO 97/06261. A preferred selection marker gene is the *A.nidulans amdS* coding sequence fused to the *A.nidulans gpdA* promoter (EP635 574B). AmdS genes from other filamentous fungus may also be used (WO 97/06261).

According to another preferred embodiment, the reporter gene encodes a selectable
5 marker for prokaryotic, yeast and/or filamentous fungal cells as described here above. Alternatively, the reporter gene encodes a reporter protein like: beta-galactosidase, beta-glucuronidase, aequorin, Green Fluorescent Protein (GFP) and variants thereof (Red, Cyan, Yellow- fluorescent protein), luciferase, lux, heme, beta-lactamase, or alkaline phosphatase. More preferably, the reporter gene is or comprises a gene encoding a fluorescent reporter
10 protein such as, but not limited to GFP, YFP, and CFP (green, yellow and cyan fluorescent proteins, respectively). Transformants will exhibit fluorescence and can be separated from non expressing clones using sensitive cell sorting technology (FACS, fluorescence activated cell sorting). An important advantage of this method is that plating of transformation mixtures is not required (which considerably increases the throughput of the
15 screening procedure) and that positive clones can arrayed directly into an MTP format for further automated processing. Alternatively, the reporter gene encodes a (trans)membrane protein or a cell wall protein to which antibodies can be raised. Transformants exhibiting said membrane protein can be separated from non expressing transformants using magnetic cell sorting technologies (MACS[®], Magnetic cell sorting, www.miltenyibiotec.com).

Control elements

The DNA construct may further comprise one or more control sequences in addition to the promoter DNA sequence, which direct the expression of the reporter gene in a suitable host cell under conditions compatible with the control sequences. Expression will
25 be understood to include any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion. One or more control sequences may be native to the reporter gene or to the host. Alternatively, one or more control sequences may be replaced with one or more control sequences foreign to the nucleic acid sequence for improving expression of
30 the reporter gene in a host cell.

The term "control sequences" is defined herein to include all components, including the promoter of the invention, which are necessary or advantageous for the expression of a coding sequence such as a reporter gene. Each control sequence may be native or foreign to the coding sequence encoding a polypeptide. Such control sequences include, but are
35 not limited to, a leader, optimal translation initiation sequences (as described in Kozak, 1991, J. Biol. Chem. 266:19867-19870), polyadenylation sequence, propeptide sequence,

signal peptide sequence, upstream activating sequence, the promoter of the invention including variants, fragments, hybrid and tandem thereof and transcription terminator. At a minimum, the control sequences include transcriptional and translational stop signals and (part of) the promoter of the invention. The control sequences may be provided with linkers
5 for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding sequence, such as a reporter gene, encoding a polypeptide.

The control sequence may be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is in operative association with the 3' terminus of the coding sequence encoding the
10 polypeptide. Any terminator, which is functional in the host cell of choice may be used in the present invention.

The skilled person would know that 5' end and 3' end of a sequence is defined for each coding sequence. For a given terminator sequence and a given coding sequence, the skilled person would know how to operatively associate them together.

15 Preferred terminators for filamentous fungal host cells are obtained from the genes for *A. oryzae* TAKA amylase, *A. niger* glucoamylase, *A. nidulans* anthranilate synthase, *A. niger* alpha-glucosidase, trpC gene, and *Fusarium oxysporum* trypsin-like protease.

Preferred terminators for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1),
20 and *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos et al, 1992, supra. The control sequence may also be a suitable leader sequence, a nontranslated region of an mRNA which is important for translation by the host cell. The leader sequence is in operative association with the 5' terminus of the nucleic acid sequence encoding the
25 polypeptide. Any leader sequence that is functional in the host cell of choice may be used in the present invention.

Preferred leaders for filamentous fungal host cells are obtained from the genes for *A. oryzae* TAKA amylase, *A. nidulans* triose phosphate isomerase and *A. niger* glucoamylase.

Suitable leaders for yeast host cells are obtained from the genes for
30 *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase and glyceraldehyde-3-phosphate dehydrogenase (ADH2 and GAP).

The control sequence may also be a polyadenylation sequence, a sequence in
35 operative association with the 3' terminus of the nucleic acid sequence and which, when Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the

genes for *A. oryzae* TAKA amylase, *A. niger* glucoamylase, *A. nidulans* anthranilate synthase, *Fusarium oxysporum* trypsin-like protease, and *A. niger* alpha-glucosidase.

Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, Molecular Cellular Biology 15: 5983-5990.

5 It may also be desirable to add regulatory sequences, which allow the regulation of the expression of the reporter gene relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the reporter gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the lac, and trp
10 operator systems of *E. coli*, and the *spac*, *xyl*, *sacB*, and *citM* systems in *B. subtilis* (reviewed in Meima et al. (2004) Expression Systems in *Bacillus*. In "Protein Expression Technologies. Current status and future trends" (F. Baneyx, ed.), Horizon Bioscience, Wymondham, Norfolk, UK, pp. 199-252). Other promoters useful for controlled expression in bacilli are those induced under nutrient starvation conditions such as the Pho regulon and
15 TnrA/GlnR system, which are under control of phosphate and nitrogen, respectively. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the TAKA alpha-amylase promoter, *A. niger* glucoamylase promoter, *A. oryzae* glucoamylase promoter, *A. tubingensis* endoxylanase (*xlnA*) promoter, *A. niger* nitrate reductase (*niaD*) promoter, *Trichoderma reesei* cellobiohydrolase promoter and the *A. nidulans* alcohol and
20 aldehyde dehydrogenase (*alcA* and *aldA*, respectively) promoters as described in US 5,503,991) may be used as regulatory sequences. Other examples of regulatory sequences are those, which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene, which is amplified in the presence of methotrexate, and the metallothionein genes, which are amplified with heavy metals. In
25 these cases, the nucleic acid sequence encoding the polypeptide would be in operative association with the regulatory sequence.

Endogenous regulatory sequence(s) present in the promoter DNA sequence of the invention may be removed, for example removal of *creA* binding sites (carbon catabolite repression as described earlier in EP673429B), change of *pacC* and *areA* (for pH and
30 nitrogen regulation). In *B. subtilis*, carbon catabolite repression (CCR) involves binding of the CcpA protein to the *cre* element. The *cre* element was used in combination with the xylose inducible *xyl* promoter for the development of an expression system, allowing for dual control of genes of interest (Bhavsar et al. (2001) *Appl Environ Microbiol* **67**: 403).

Preferably, the DNA construct comprises a promoter DNA sequence from the
35 invention, a reporter gene in operative association with said promoter DNA sequence and translational control sequences such as:

- one translational termination sequence orientated in 5' towards 3' direction selected from the following list of sequences: TAAG, TAGA and TAAA, preferably TAAA, and/or

- one translational initiator coding sequence orientated in 5' towards 3' direction selected from the following list of sequences: GCTACCCCC; GCTACCTCC; GCTACCCTC; GCTACCTTC; GCTCCCCCC; GCTCCCTCC; GCTCCCCTC; GCTCCCTTC; GCTGCCCCC; GCTGCCTCC; GCTGCCCTC; GCTGCCTTC; GCTTCCCCC; GCTTCCTCC; GCTTCCCTC; and GCTTCCTTC, preferably GCT TCC TTC, and/or

- one transcriptional initiator sequence selected from the following list of sequences: 5'-mwChkyCAAA-3'; 5'-mwChkyCACA-3' or 5'-mwChkyCAAG-3', using ambiguity codes for nucleotides: m (A/C); w (A/T); y (C/T); k (G/T); h (A/C/T), preferably 5'-CACCGTCAAA-3' or 5'-CGCAGTCAAG-3'.

In the context of this invention, the term "translational initiator coding sequence" is defined as the nine nucleotides immediately downstream of the initiator or start codon of the open reading frame of a DNA coding sequence. The initiator or start codon encodes for the AA methionine. The initiator codon is typically ATG, but may also be any functional start codon such as GTG.

In the context of this invention, the term "translational termination sequence" is defined as the three or four nucleotides starting from the translational stop codon at the 3' end of the open reading frame or nucleotide coding sequence and oriented in 5' towards 3' direction.

In the context of this invention, the term "translational initiator sequence" is defined as the ten nucleotides immediately upstream of the initiator or start codon of the open reading frame of a DNA sequence coding for a polypeptide. The initiator or start codon encodes for the AA methionine. The initiator codon is typically ATG, but may also be any functional start codon such as GTG. It is well known in the art that uracil, U, replaces the deoxynucleotide thymine, T, in RNA.

Preferred control sequences for prokaryotes have already been described in EP 284 126 A.

The present invention also relates to recombinant expression vectors comprising a promoter of the present invention operatively associated with a reporter gene, and transcriptional and translational stop signals. The various reporter gene and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the promoter and/or reporter gene at such sites. Alternatively, fusion of

reporter gene and promoter can be done by e.g. sequence overlap extension using PCR (SOE-PCR), as described in Gene. 1989 Apr 15;77(1):51-9. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR "Site-directed mutagenesis by overlap extension using the polymerase chain reaction") or by cloning using the Gateway™ cloning system (Invitrogen).

5 The recombinant expression vector may be any vector capable of transforming a host cell. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids. The vectors may be integrative or autonomously replicating vectors.

10 The vector may be an autonomously replicating vector, i.e., a vector, which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The autonomously replicating vector may contain any means for assuring self-replication. The vector may be one which, when introduced into the host cell,
15 is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used. An example of an autonomously maintained cloning vector is a cloning vector comprising the AMA1-sequence. AMA1 is a 6.0-kb genomic DNA
20 fragment isolated from *A. nidulans*, which is capable of Autonomous Maintenance in *Aspergillus* (see e.g. Aleksenko and Clutterbuck (1997), Fungal Genet. Biol. **21**: 373-397). Examples of autonomously replicating vectors useful for expression in Gam-positive bacteria such as *B. subtilis* and *B. amyloliquefaciens* are the staphylococcal vector pUB110 (McKenzie *et al.* (1986) *Plasmid* **15**: 93) and the endogenous *B. subtilis* plasmids pTA1015,
25 pTA1040 and pTA1060 and vectors derived thereof (e.g. pHB201; Bron *et al.* (1998) *J Biotechnol* 64: 3).

The vectors of the present invention preferably contain (an) element(s) that permits stable integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

30 For integration into the host cell genome, the vector may rely on the promoter sequence and/or reporter gene sequence or any other element of the vector for stable integration of the vector into the genome by homologous or non-homologous recombination. The vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid
35 sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise

location, the integrational elements should preferably contain a sufficient number of nucleic acids, such as 20 to 50, preferably, 55 to 95, preferably 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, more preferably 800 to 1,500 base pairs, and most preferably at least 2kb, which are homologous to a DNA sequence in a predetermined target locus in the genome of the host cell. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences. In order to promote targeted integration, the cloning vector is preferably linearized prior to transformation of the host cell. Linearization is preferably performed such that at least one but preferably either end of the cloning vector is flanked by sequences homologous to the target locus.

Preferably, the integrational elements in the expression vector, which are homologous to the target locus are derived from a highly expressed locus meaning that they are derived from a gene, which is capable of high expression level in the host cell. A gene capable of high expression level, i.e. a highly expressed gene, is herein defined as a gene whose mRNA can make up at least 0.5% (w/w) of the total cellular mRNA, e.g. under induced conditions, or alternatively, a gene whose gene product can make up at least 1% (w/w) of the total cellular protein, or, in case of a secreted gene product, can be secreted to a level of at least 0.1 g/l (as described in EP 357 127 B1). A number of preferred highly expressed fungal genes are given by way of example: the amylase, glucoamylase, alcohol dehydrogenase, xylanase, glyceraldehyde-phosphate dehydrogenase or cellobiohydrolase genes from *Aspergilli* or *Trichoderma*. Most preferred highly expressed genes for these purposes are a glucoamylase gene, preferably an *A. niger* glucoamylase gene, an *A. oryzae* TKA-amylase gene, an *A. nidulans gpdA* gene, the locus of SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28 and SEQ ID NO:29, preferably the *A. niger* locus of SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28 and SEQ ID NO:29 or a *Trichoderma reesei* cellobiohydrolase gene.

Alternatively, the vector may be integrated into the genome of the host cell by non-homologous recombination.

Alternatively and/or in combination with the embodiments described above, the DNA construct comprising the reporter gene in operative association with the promoter of the invention may be deleted from the host cell after the cell has been screened for production, preferably secretion, of a protein of interest in step (e) of the invention. Step (e) is described later. This enables the host cell to be used directly for the production of the protein of interest, without said host cell comprising a reporter gene and/or selectable marker gene. To enable deletion of the DNA construct of the invention, the DNA construct preferably

comprises both the reporter gene and a bi-directional selectable marker gene between DNA repeats. The DNA repeats allow deletion of the DNA construct by intra chromosomal homologous recombination. This method is analogous to the "MARKER-GENE FREE" approach as described in EP 0 635 574 B1.

5 For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6. The origin of replication may be one having a mutation which makes its functioning temperature-sensitive in the host cell (see, e.g., Ehrlich, 1978, Proceedings of the National Academy of Sciences USA 75:1433).

10 More than one copy of the DNA construct of the present invention may be inserted into the host cell. This can be done, preferably by integrating into its genome copies of the DNA sequence, more preferably by targeting the integration of the DNA sequence at a highly expressed locus, preferably at a glucoamylase locus or at the locus of SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28 and SEQ ID NO:29. Alternatively, this can be done by including an amplifiable selectable marker gene with the nucleic acid sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleic acid sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent. To increase even more the number of copies of the DNA sequence to be over expressed the technique of gene conversion as described in WO98/46772 may be used.

20 The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, e.g., Sambrook et al., 1989, supra).

Host cell

30 The present invention further relates to recombinant host cells transformed with the DNA construct defined in the former section comprising a promoter DNA sequence of the present invention in operative association with a reporter gene. Preferably, the transformed host cell further comprises an additional DNA construct comprising a DNA sequence comprising a coding sequence originating from a DNA library from an organism suspected of being capable of producing one or more proteins with properties of interest.

35 Such cells can be advantageously used in an expression cloning method as described in the next section. An expression vector comprising a promoter DNA sequence of the present invention in operative association with a reporter gene is introduced into a

host cell. The host cell is further transformed with a DNA construct or expression vector comprising a DNA sequence comprising a coding sequence originating from a DNA library. Both DNA construct or expression vectors are either maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The choice of a host cell will to a large extent depend upon the organism wherein the DNA library has been made and/or upon the origin of the promoter DNA sequence used.

The host cell may be any microorganism. Preferably, the host cell is the same organism as the one wherein the DNA library has been prepared and/or as the one the promoter of the invention originates from as defined earlier in the specification. According to a preferred embodiment, the host cell is a prokaryote, or an eukaryote. According to a more preferred embodiment, the host cell useful in the methods of the present invention is a prokaryote, more preferably the prokaryote defined as source of promoter in the section promoter. More preferably, the prokaryotic host cell is a *Bacillus* host cell, preferably of the species *Bacillus Subtilis*, *Bacillus licheniformis* and/or *Bacillus amyloliquefaciens*.

According to another more preferred embodiment, the host cell useful in the methods of the present invention is a fungus.

The host cell may be a wild type filamentous fungus host cell or a variant, a mutant or a genetically modified filamentous fungus host cell. In a preferred embodiment of the invention the host cell is a protease deficient or protease minus strain. This may be the protease deficient strain *Aspergillus oryzae* JaL 125 having the alkaline protease gene named "alp" deleted (described in WO 97/35956 or EP 429 490), or the tripeptidyl-aminopeptidases (TPAP) deficient strain of *A. niger*, disclosed in WO 96/14404. Further, also host cell with reduced production of the transcriptional activator (prtT) as described in WO 01/68864 is contemplated according to the invention. Another specifically contemplated host strain is the *Aspergillus oryzae* BECh2, where the three TKA amylase genes present in the parent strain IF04177 have been inactivated. In addition, two proteases, the alkaline protease and neutral metalloprotease 11 have been destroyed by gene disruption. The ability to form the metabolites cyclopiazonic acid and kojic acid has been destroyed by mutation. BECh2 is described in WO 00/39322 and is derived from JaL228 (described in WO 98/12300), which again was a mutant of IF04177 disclosed in US 5,766, 912 as A1560.

Optionally, the filamentous fungal host cell may comprise an elevated unfolded protein response (UPR) compared to the wild type cell to enhance production abilities of a polypeptide of interest. UPR may be increased by techniques described in US2004/0186070A1 and/or US2001/0034045A1 and/or WO01/72783A2 and/or WO2005/123763. More specifically, the protein level of HAC1 and/or IRE1 and/or PTC2 has

been modulated, and/or the SEC61 protein has been engineered in order to obtain a host cell having an elevated UPR.

Alternatively, or in combination with an elevated UPR, the host cell may be genetically modified to obtain a phenotype displaying lower protease expression and/or protease secretion compared to the wild-type cell in order to enhance production abilities of a polypeptide of interest. Such phenotype may be obtained by deletion and/or modification and/or inactivation of a transcriptional regulator of expression of proteases. Such a transcriptional regulator is e.g. prtT. Lowering expression of proteases by modulation of prtT may be performed by techniques described in US2004/0191864A1 and in EP2005/055145.

Alternatively, or in combination with an elevated UPR and/or a phenotype displaying lower protease expression and/or protease secretion, the host cell may display an oxalate deficient phenotype in order to enhance the yield of production of a polypeptide of interest. An oxalate deficient phenotype may be obtained by techniques described in WO2004/070022A2 and/or WO2000/50576.

Alternatively, or in combination with an elevated UPR and/or a phenotype displaying lower protease expression and/or protease secretion and/or oxalate deficiency, the host cell may display a combination of phenotypic differences compared to the wild cell to enhance the yield of production of the polypeptide of interest. These differences may include, but are not limited to, lowered expression of glucoamylase and/or neutral alpha-amylase A and/or neutral alpha-amylase B, alpha-1, 6transglucosidase, protease, and oxalic acid hydrolase. Said phenotypic differences displayed by the host cell may be obtained by genetic modification according to the techniques described in US2004/0191864A1.

Alternatively, or in combination with phenotypes described here above, the efficiency of targeted integration of a nucleic acid construct into the genome of the host cell by homologous recombination, i.e. integration in a predetermined target locus, may preferably be increased by augmented homologous recombination abilities of the host cell. Such phenotype of the cell preferably involves a deficient *hdfA* or *hdfB* gene as described in WO2005/095624. WO2005/095624 discloses a preferred method to obtain a filamentous fungal cell comprising increased efficiency of targeted integration.

More preferably, the host cell of the invention is selected from the following list: a *Bacillus*, a yeast and a filamentous fungus, preferably an *Aspergillus*, *Penicillium* or *Trichoderma* species. Even more preferably, the *Aspergillus* host cell is an *Aspergillus niger* or *Aspergillus sojae* or *Aspergillus oryzae* species.

The present invention also relates to recombinant host cells, comprising more than one promoter DNA sequence of the present invention, each promoter being in operative association with a reporter gene. Such host cells may be advantageously used in an

expression cloning method as described in the coming section. Bacterial cells may be transformed by a variety of methods including chemically induced competence (e.g. CaCl_2) or electroporation (Sambrook *et al.* (1989) "Molecular Cloning: a laboratory manual", Cold Spring Harbor Laboratories, Cold Spring Harbor, New York), protoplast transformation (Chang and Cohen (1979) *Mol Gen Genet* **168**: 111) and, in the case of some *Bacillus* species, natural competence (Spizizen (1958) *Proc Natl Acad Sci USA* **44**: 1072). Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* host cells are described in EP 238 023 and Yelton *et al.*, 1984, Proceedings of the National Academy of Sciences USA 81: 1470-1474. Suitable procedures for transformation of *Aspergillus* and other filamentous fungal host cells using *Agrobacterium tumefaciens* are described in e.g. Nat Biotechnol. 1998 Sep;16(9):839-42. Erratum in: Nat Biotechnol 1998 Nov;16(11):1074. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. de Groot MJ, Bundock P, Hooykaas PJ, Beijersbergen AG. Unilever Research Laboratory Vlaardingen, The Netherlands. Suitable methods for transforming *Fusarium* species are described by Malardier *et al.*, 1989, Gene 78: 147-156 and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, In Abelson, J. N. and Simon, M. I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito *et al.*, 1983, Journal of Bacteriology 153: 163; and Hinnen *et al.*, 1978, Proceedings of the National Academy of Sciences USA 75: 1920. Transformation of prokaryotes may be performed as described in EP 284 126A.

Method of expression cloning

According to a further aspect, the present invention provides a method for isolating a DNA sequence comprising a DNA sequence coding for a protein of interest in a host cell, said method comprises the steps of:

- (a) preparing a first DNA construct comprising a promoter DNA sequence operatively associated with a reporter gene conferring a selectable trait; said promoter DNA sequence being induced when the DNA construct is present in the host cell and when a protein of interest, preferably a secreted protein of interest, is produced by the host cell;
- (b) preparing a second DNA construct comprising a DNA sequence comprising a DNA sequence coding for a protein of interest originating from a DNA library from an organism suspected of being capable of producing one or more proteins of interest;

- (c) transforming a host cell with both DNA constructs prepared in (a) and (b);
(d) culturing all the transformed host cells obtained in (c) under conditions conducive to the production of the proteins of interest as present in the DNA library; and
5 (e) screening for transformed host cells producing a protein of interest by analysis of the proteins produced in (d).

Step (a)

A first DNA construct is prepared, comprising a promoter DNA sequence operatively associated with a reporter gene conferring a selectable trait; said promoter DNA sequence
10 being induced when the DNA construct is present in the host cell and when a protein of interest, preferably a secreted protein of interest, is produced by the host cell.

The promoter present in the first DNA construct is a promoter which is induced when a protein of interest, preferably a secreted protein of interest, is produced by a host cell.

In the present invention the term "induced" is defined as an increase in promoter
15 activity of a promoter of the invention when a protein of interest is (over)produced, preferably secreted, by a host cell, compared to the activity of the promoter when the protein of interest is not (over)produced in a corresponding control host cell under the same culture conditions. Preferably, when a protein of interest is (over)produced, preferably secreted, in a
20 host cell, promoter activity is at least about 1.5-fold increased, more preferably at least about two-fold, more preferably at least about three-fold, more preferably at least about four-fold, more preferably at least about five-fold, even more preferably at least about six-fold, even more preferably at least about eight-fold, even more preferably at least about ten-fold, even more preferably at least about twenty-fold, even more preferably at least about 50-fold and most preferably at least about 100-fold increased, compared to a corresponding control
25 host cell wherein the protein of interest is not (over)produced under the same culture conditions. More preferably, the increase in promoter activity is infinite, i.e. no promoter activity is detected or the promoter is inactive when the protein of interest is not (over)produced, whereas the promoter is induced when the protein of interest is (over)produced. Promoter activity is preferably determined as defined previously. Most
30 preferably, promoter activity is determined by mRNA analysis using Northern blot.

Such a promoter can be identified by comparing gene expression profiles of cells producing the protein of interest with that of a control strain not producing said protein. Preferably, such expression profiles are compared using well-established DNA micro-array analyses. Alternatively, expression profiles can be compared using other
35 methods known to the skilled person such as Northern blot or quantitative PCR analysis. A most preferred method to compare expression profiles is performed by first comparing

expression profiles by DNA micro array analyses, followed by subsequent corroboration of the micro array results by Northern Blot. According to a preferred embodiment, the promoter is the promoter DNA sequence of the invention described under the section promoter.

5 According to another preferred embodiment, when the host cell is a *Bacillus* cell, the promoter is derived from the following genomic DNA sequences as listed: *htrA* (SEQ ID NO:1) or *htrB* (SEQ ID NO:2). Preferably, the promoter DNA sequence is derived from the non coding part of one of these genomic DNA sequences situated upstream of the start codon or part of said genomic sequences situated upstream. More preferably, the promoter
10 used is the DNA sequence of either SEQ ID NO:1 or SEQ ID NO:2.

According to another preferred embodiment, when the host cell is a filamentous fungal cell, the promoter is derived from the following genomic DNA sequences as listed: SEQ ID NO:20, SEQ NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, or SEQ ID NO: 24. Preferably, the promoter DNA sequence is derived from the non coding part of one of these genomic DNA
15 sequences situated upstream of the start codon or part of said genomic sequences situated upstream. More preferably, the promoter used is the DNA sequence, or derivative thereof, of either SEQ ID NO:15, SEQ NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, or SEQ ID NO: 19.

The first DNA construct prepared in step (a) has already been described in the corresponding section "DNA construct" and preferably comprises a reporter gene conferring
20 a selectable trait as described in the section "DNA construct".

Step (b)

The DNA construct prepared in step (b) comprises a DNA sequence comprising a DNA sequence coding for a protein of interest originating from a DNA library from an
25 organism suspected of being capable of producing one or more proteins of interest. The organism suspected of producing one or more proteins of interest usually is a prokaryote or an eukaryote. Examples of preferred prokaryotes and eukaryotes have already been defined earlier in the specification in connection with the origin of the promoter of the invention. According to one preferred embodiment, the organism is a *Bacillus*, more preferably a
30 *Bacillus* species as defined in the section promoter. According to another preferred embodiment, the organism is an eukaryote, more preferably a fungus, of which most preferably a filamentous fungus.

The DNA construct prepared in step (b), preferably comprises regulatory sequences such as defined in the sections "DNA construct" and "Control elements".

35 In the method according to the invention, the library of DNA fragments from an organism suspected of producing one or more proteins of interest can be a genomic library

or a cDNA library. However, in the case of eukaryotic donors, preferably a cDNA library is used so as to avoid problems with recognition of promoters or splice signals in the fungal host organism. The cDNA library is preferably prepared from mRNA isolated from the source organism when grown under conditions conducive to the expression of the proteins of interest.

The method according to the invention can be applied to the isolation of DNA sequences coding for any protein or polypeptide of interest if there is an assay available for detection of the protein when expressed by the host cell.

Step (c)

The identity of the host cell to be transformed with both DNA constructs has already been described under the section host cell. Methods of transformation have already been described in the same section. According to one embodiment, the host cell is simultaneously transformed with the constructs prepared in (a) and (b). Alternatively and according to another embodiment, the host cell is first transformed with the DNA construct prepared in step (a) and consecutively is transformed with the DNA construct prepared in step (b).

According to a preferred embodiment, the host cell is a prokaryote, preferably a *Bacillus* cell, more preferably a *Bacillus subtilis*, even more preferably a *Bacillus subtilis* deficient in the *htrA* gene, wherein at least part of the coding regions of the *htrA* gene have been deleted and/or replaced and/or wherein the non coding regions of the *htrA* gene are still present in the genome of said *Bacillus* cell, most preferably the *Bacillus subtilis* BV2003 is used (Hyrylainen et al. (2001) *Mol Microbiol* **41**: 1159).

Alternatively or in combination with former preferred embodiment, the host cell is a *Bacillus* cell and the reporter gene encodes a fluorescence reporter protein as defined in the section DNA construct, most preferably GFP.

Alternatively or in combination with former preferred embodiment, when the host cell is a *Bacillus* cell, the promoter is derived from the genomic DNA sequences as presented in the following list: *htrA* (SEQ ID NO:1) or *htrB* (SEQ ID NO:2). Preferably, the promoter sequence is derived from the non coding part of one of these two genomic DNA sequences situated upstream of the start codon or part of said sequence situated upstream. More preferably, the promoter used is one derived from either SEQ ID NO:1 or SEQ ID NO:2. According to a most preferred embodiment, the promoter used and the reporter gene used are either GFP in operative association with SEQ ID NO: 1 or GFP in operative association with SEQ ID NO: 2.

According to another preferred embodiment, the organism is an eukaryote, more preferably a fungus, of which most preferably a filamentous fungus. Preferably, the filamentous fungus is an *Aspergillus*, *Penicillium* or *Trichoderma* species. Even more preferably, the *Aspergillus* host cell is an *Aspergillus niger* or *Aspergillus sojae* or *Aspergillus oryzae* species. A most preferred host cell is *A.niger*, preferably CBS513.88 or derivative thereof.

Alternatively or in combination with former preferred embodiment, the host cell is a filamentous fungal cell and the reporter gene encodes a fluorescence reporter protein as defined in the section DNA construct, most preferably GFP. Alternatively, the reporter gene encodes a selectable marker. More preferably, the reporter gene is a hybrid reporter gene of a fluorescent protein and a selectable marker. Most preferably, the hybrid reporter gene is a comprises GFP and the phleomycin resistance gene (BLE) of *Streptoalloteichus hindustanis* (Drocourt *et al.*, NAR, 1990).

Alternatively or in combination with the former preferred embodiment, when the host cell is a filamentous fungal cell, the promoter is derived from the genomic DNA sequences as presented in the following list: SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24. Preferably, the promoter sequence is derived from the non coding part of one of the listed genomic DNA sequences situated upstream of the start codon or part of said sequence situated upstream. More preferably, the promoter used is one of those used in examples 3 to 7 and derived from either SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24. According to a most preferred embodiment, the promoter used and the hybrid reporter gene used are either GFP-BLE (SEQ ID NO: 61) in operative association with promoter SEQ ID NO: 15, GFP-BLE in operative association with promoter SEQ ID NO: 16, GFP-BLE in operative association with promoter SEQ ID NO: 17, GFP-BLE in operative association with promoter SEQ ID NO: 18, or GFP-BLE in operative association with promoter SEQ ID NO: 19.

Step (d)

After transformation of the host cells with both DNA constructs, all the obtained transformed host cells clones are cultured under conditions conducive to the production of the proteins of interest present in the DNA library. Depending on the assay required for detection of the protein of interest the transformed clones are propagated and stored as colonies on solid media such as agar plates or in liquid media, whereby the individual library clones are grown, stored and/or assayed in the wells of the microtiter plates.

The skilled person will understand that the usual adaptations to cloning methods known in the art can equally be applied to the method of the present invention. The

adaptations include but are not limited to e.g. screening of pools of library clones, screening the same library for a number of different proteins of interest, as well as rescreening, reisolation and recloning of positive clones to ensure more accurate results.

A variety of methods are available to the skilled person for isolation of the DNA sequence encoding the protein of interest from the transformed host cell identified in the screening method, and for subsequent characterization of the isolated DNA sequence.

Step (e)

Subsequent to step (d), the transformed host cells are screened for production, preferably secretion, of a protein of interest by analysis of the proteins produced. During step (e), one screens for transformed host cells producing, preferably secreting a protein of interest, by monitoring production-dependent, preferably secretion-dependent reporter expression induced via the promoter DNA sequence present in the first DNA construct.

According to a preferred embodiment, when the host cell is a *Bacillus* cell, screening to identify transformed host cells producing, preferably secreting, a protein of interest by monitoring production-dependent, preferably secretion-dependent, reporter expression induced by the promoter DNA sequence present in the first DNA construct, is preferably performed using fluorescent based cell analysis assays, e.g. fluorescent activated cell scanning, fluorescent activated cell sorting (FACS) or fluorimetric analysis. More preferably, the host cell is a *Bacillus* cell comprising a first DNA construct, said DNA construct comprising a promoter such as SEQ ID NO: 1 or derivative thereof or SEQ ID NO: 2 or derivative thereof, said DNA construct further comprising a reporter gene encoding a fluorescent reporter and screening is performed using fluorescence based cell analysis assays. Even more preferably, the host cell is a *Bacillus* cell and the first DNA construct comprises:

- the promoter DNA sequence SEQ ID NO: 1 or derivative thereof, in operative association with GFP, or
- the promoter DNA sequence SEQ ID NO: 2 or derivative thereof, in operative association with GFP

and screening is performed using FACS. Alternatively, or in combination with the former embodiment, screening may be performed using e.g. colorimetric assays or using selective culture conditions. The skilled person will understand that depending on the reporter gene used, the selective culture conditions should be chosen accordingly. Examples of selective culture conditions are, but are not limited to, the use of biocides, the use of antibiotics, limitation of at least one nutrient, prototrophy.

According to another preferred embodiment, when the host cell is a filamentous fungal cell, screening to identify transformed host cells producing, preferably secreting, a protein of interest by monitoring production-dependent, preferably secretion-dependent, reporter expression induced by the promoter DNA sequence present in the first DNA construct of the invention, is preferably performed by cultivation under selective culture conditions. The skilled person will understand that depending on the reporter gene used, the selective culture conditions should be chosen accordingly. Examples of selective culture conditions are, but are not limited to, the use of biocides, the use of antibiotics, limitation of at least one nutrient, prototrophy. Alternatively, or in combination with the former embodiment, screening may be performed using e.g. colorimetric, fluorimetric (like FACS) or enzyme activity based assays. According to a more preferred embodiment, the host cell is a filamentous fungal cell comprising a first DNA construct of the invention and screening is performed using selective culture conditions. More preferably, the host cell is a filamentous fungal cell comprising a first DNA construct of the invention, said DNA construct comprising a promoter selected from the list of: SEQ ID NO:15 or derivative thereof, SEQ ID NO:16 or derivative thereof, SEQ ID NO:17 or derivative thereof, SEQ ID NO:18 or derivative thereof, SEQ ID NO:19 or derivative thereof, said DNA construct further comprising a reporter gene encoding a selectable marker and screening is performed using selective culture conditions. Even more preferably, the host cell is a filamentous fungal cell comprising a first DNA construct of the invention, said DNA construct comprising a promoter selected from the list of: SEQ ID NO:15 or derivative thereof, SEQ ID NO:16 or derivative thereof, SEQ ID NO:17 or derivative thereof, SEQ ID NO:18 or derivative thereof, SEQ ID NO:19 or derivative thereof, said DNA construct further comprising a hybrid reporter gene such as SEQ ID NO: 61 and screening is performed using selective culture conditions.

According to another more preferred embodiment, the host cell is an *A.niger* cell comprising a first DNA construct of the invention and screening is performed using selective culture conditions. More preferably, the host cell is an *A.niger* cell comprising a first DNA construct of the invention, said DNA construct comprising a promoter selected from the list of: SEQ ID NO:15 or derivative thereof, SEQ ID NO:16 or derivative thereof, SEQ ID NO:17 or derivative thereof, SEQ ID NO:18 or derivative thereof, SEQ ID NO:19 or derivative thereof, said DNA construct further comprising a reporter gene encoding a selectable marker and screening is performed using selective culture conditions. Even more preferably, the host cell is an *A.niger* cell comprising a first DNA construct of the invention, said DNA construct comprising a promoter selected from the list of: SEQ ID NO:15 or derivative thereof, SEQ ID NO:16 or derivative thereof, SEQ ID NO:17 or derivative thereof,

SEQ ID NO:18 or derivative thereof, SEQ ID NO:19 or derivative thereof, said DNA construct further comprising a hybrid reporter gene such as SEQ ID NO:61 and screening is performed using selective culture conditions. Even more preferably, the host cell is an *A.niger* cell and the first DNA construct comprises:

- 5 - hybrid reporter gene GFP-BLE (SEQ ID NO: 61) in operative association with promoter SEQ ID NO: 15, or
- GFP-BLE in operative association with promoter SEQ ID NO: 16, or derivative thereof, or
- GFP-BLE in operative association with promoter SEQ ID NO: 17, or derivative thereof, or
- 10 - GFP-BLE in operative association with promoter SEQ ID NO: 18, or derivative thereof, or
- GFP-BLE in operative association with promoter SEQ ID NO: 19, or derivative thereof, and screening is performed using selective culture conditions comprising an antibiotic, preferably phleomycin as selective agent.

The DNA sequences isolated by the screening method of the invention as described above are used to produce, or to improve the production of, a protein of interest encoded by the DNA sequence. Advantageously, the transformed host cell as isolated in the above described screening method is used directly in a process for the production of the protein of interest by culturing the transformed host cell under conditions conducive to the production, preferably secretion of the protein of interest and, optionally, recovering the protein.

According to a preferred embodiment, the host cell is enabled to be used directly for the production of the protein of interest, without said host cell comprising a reporter gene and/or selectable marker gene. The DNA construct comprising the reporter gene in operative association with the promoter of the invention is therefore deleted from the host cell after the cell has been screened for production, preferably secretion, of a protein of interest. To enable deletion of the DNA construct of the invention, the DNA construct preferably comprises both the reporter gene and a bi-directional selectable marker gene between DNA repeats. The DNA repeats allow deletion of the DNA construct by intra chromosomal homologous recombination. This method is analogous to the "MARKER-GENE FREE" approach as described in EP 0 635 574 B1.

Often the initial transformed host cell isolated in the screening method of the invention will have an expression level which is satisfactory for screening purposes but which can be significantly improved for economic production purposes. To this end, the DNA sequence is inserted into an expression vector which is subsequently used to transform a suitable host cell. In the expression vector the DNA sequence is operably linked

to appropriate expression signals, such as a promoter, optionally a signal sequence and a terminator, which are capable of directing the expression of the protein in the host organism as defined in the sections "DNA construct" and "control elements". A suitable host cell for the production of the protein may be either a prokaryotic or eukaryotic cell, preferably a eubacterium, yeast or filamentous fungus. Preferred bacterial host cells are selected from the genus *Bacillus* as defined in the section promoter. According to another preferred embodiment, the host cell is a yeast or a filamentous fungus as already defined in the section "host cell". Preferred yeast host cells are selected from the group consisting the genera *Saccharomyces*, *Kluyveromyces*, *Yarrowia*, *Pichia*, and *Hansenula*. Preferred filamentous fungal host cells are selected from the same genera listed above as preferred host cells for the screening method. More preferably, the filamentous fungus is an *Aspergillus*, *Penicillium* or *Trichoderma* species. Even more preferably, the *Aspergillus* host cell is an *Aspergillus niger* or *Aspergillus sojae* or *Aspergillus oryzae* species. A most preferred host cell is *A.niger*, preferably CBS513.88 or derivative thereof.

The suitable host cell is transformed with the expression vector by any of the various protocols available to the skilled person. The transformed host cell is subsequently used in a process for producing the protein of interest by culturing the transformed host cell under conditions conducive to the expression of the DNA sequence encoding the protein, and optionally recovering the protein.

Preferably, the protein of interest is an enzyme.

Accordingly, the invention relates to a use of the promoter of the invention in an expression cloning method.

The invention further relates to other use of the promoter of the invention. Preferably, the promoter of the invention may be operatively associated with any coding sequence in a DNA construct or expression vector all defined in former sections. Such DNA construct or expression vector may be introduced into any host cell as defined in "host cell" section and used for expression of a given valuable compound. Examples of such valuable compound are, but are not limited to, a metabolite or a polypeptide. Preferably, the coding sequence operatively associated with the promoter of the invention encodes the valuable compound. Alternatively, and according to another preferred embodiment of the invention, the coding sequence operatively associated with the promoter of the invention is involved in the production of the valuable compound.

Alternatively and according to another preferred use of the promoter of the invention, the promoter of the present invention is incorporated into an inactivation or replacement construct, transformed into a given host cell and used to inactivate or replace a target gene.

The present invention is further illustrated by the following examples.

Examples

Materials and methods

5 General procedures

Standard molecular cloning techniques such as DNA isolation, gel electrophoresis, enzymatic restriction modifications of nucleic acids, Northern analyses, *E. coli* transformation, etc, were performed as described by Sambrook et al. (2001) "Molecular Cloning: a laboratory manual, 3^d edition", Cold Spring Harbor Laboratories, Cold Spring Harbor, New York and Innis et al. (1990) "PCR protocols, a guide to methods and applications" Academic Press, San Diego. Synthetic oligo deoxynucleotides were obtained from Invitrogen (Breda, The Netherlands). Room temperature is 20 degrees Celsius \pm 2 degrees Celsius.

15 Transformation of *Aspergillus niger*.

Transformation of *A. niger* was performed according to the method described by Tilburn, J. et al. (1983) Gene 26, 205-221 and Kelly, J. & Hynes, M. (1985) EMBO J., 4, 475-479 with the following modifications:

Spores were germinated and cultivated for 16 hours at 30 degrees Celsius in a shake flask placed in a rotary shaker at 300 rpm in *Aspergillus* minimal medium (100ml). *Aspergillus* minimal medium contains per litre: 6 g NaNO₃, 0.52 g KCl, 1.52 g KH₂PO₄, 1.12 ml 4 M KOH, 0.52 g MgSO₄.7H₂O, 10 g glucose, 1 g casaminoacids, 22 mg ZnSO₄.7H₂O, 11 mg H₃BO₃, 5 mg FeSO₄.7H₂O, 1.7 mg CoCl₂.6H₂O, 1.6 mg CuSO₄.5H₂O, 5 mg MnCl₂.2H₂O, 1.5 mg Na₂MoO₄.2H₂O, 50 mg EDTA, 2 mg riboflavin, 2 mg thiamine-HCl, 2 mg nicotinamide, 1 mg pyridoxine-HCL, 0.2 mg panthotenic acid, 4 g biotin, 10 ml Penicillin (5000 IU/ml) Streptomycin (5000 UG/ml) solution (Gibco).

- Novozym 234TM (Novo Industries) instead of helicase was used for the preparation of protoplasts;
- after protoplast formation (60-90 minutes), KC buffer (0.8 M KCl, 9.5 mM citric acid, pH 6.2) was added to a final volume of 45 ml, the protoplast suspension was centrifuged for 10 minutes at 3000 rpm at 4 degrees Celsius in a swinging-bucket rotor. The protoplasts were resuspended in 20 ml KC buffer and subsequently 25 ml of STC buffer (1.2 M sorbitol, 10 mM Tris-HCl pH 7.5, 50 mM CaCl₂) was added. The protoplast suspension was centrifuged for 10 minutes at 3000 rpm at 4 degrees Celsius in a swinging-bucket rotor, washed in STC-buffer and resuspended in STC-buffer at a concentration of 10E8 protoplasts/ml;

- to 200 microliter of the protoplast suspension, the DNA fragment, dissolved in 10 microliter TE buffer (10 mM Tris-HCl pH 7.5, 0.1 mM EDTA) and 100 microliter of PEG solution (20% PEG 4000 (Merck), 0.8 M sorbitol, 10 mM Tris-HCl pH 7.5, 50 mM CaCl_2) was added;

5 - after incubation of the DNA-protoplast suspension for 10 minutes at room temperature, 1.5 ml PEG solution (60% PEG 4000 (Merck), 10 mM Tris-HCl pH 7.5, 50 mM CaCl_2) was added slowly, with repeated mixing of the tubes. After incubation for 20 minutes at room temperature, suspensions were diluted with 5 ml 1.2 M sorbitol, mixed by inversion and centrifuged for 10 minutes at 4000 rpm at room temperature. The protoplasts were
10 resuspended gently in 1 ml 1.2 M sorbitol and plated onto solid selective regeneration medium consisting of either *Aspergillus* minimal medium without riboflavin, thiamine.HCl, nicotinamide, pyridoxine, panthotenic acid, biotin, casaminoacids and glucose. In case of acetamide selection the medium contained 10 mM acetamide as the sole nitrogen source and 1 M sucrose as osmoticum and C-source. Alternatively, protoplasts were plated onto
15 PDA (Potato Dextrose Agar, Oxoid) supplemented with 1-50 microgram/ml phleomycin and 1M sucrose as osmoticum. Regeneration plates were solidified using 2% agar (agar No.1, Oxoid L11). After incubation for 6-10 days at 30 degrees Celsius, conidiospores of transformants were transferred to plates consisting of *Aspergillus* selective medium (minimal medium containing acetamide as sole nitrogen source in the case of acetamide
20 selection or PDA supplemented with 1-50 microgram/ml phleomycin in the case of phleomycin selection) with 2% glucose and 1.5% agarose (Invitrogen) and incubated for 5-10 days at 30 degrees Celsius. Single transformants were isolated and this selective purification step was repeated once upon which purified transformants were stored.

25 ***Aspergillus niger* microtiterplate fermentation and sampling**

To obtain spores of *A. niger*, mycelium was transferred to microtiterplates filled with 150 microliter solid PDA medium (Potato Dextrose Agar, Oxoid), prepared according to the supplier's instructions. After growth for 3-7 days at 30 degrees Celsius (static) spores had formed. Subsequently 100 microliter liquid fermentation medium (70 g/l glucose. H_2O , 25 g/l
30 peptone (from casein), 12.5 g/l yeast extract (Difco), 2 g/l K_2SO_4 , 1 g/l KH_2PO_4 , 0.5 g/l $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.03 g/l ZnCl_2 , 0.02 g/l CaCl_2 , 9 mg/l $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 3 mg/l $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, adjusted to pH 5.6 with 4 N H_2SO_4), was added to each well. The fermentation medium was pipetted up and down 20 times to isolate the spores in the medium. The inoculated fermentation medium was transferred to another microtiterplate and incubated at 34
35 degrees Celsius, 550 rpm and 80% humidity for 6 days in an orbital microtiterplate shaker

(Infors HT). On day 6 supernatants were collected and used for detection of secreted proteins.

Mass spectrometry analysis

Supernatants were ultra-filtrated by centrifugal force on an Ultracel filtration plate (Millipore, Billerica, MA, USA). Ultra-filtration was performed according to the protocol provided by the manufacturer. To the ultra-filtrated supernatants 10 microliter 200mM ammoniumbicarbonate, pH 7,8 (Riedel-de Haen AG, Seelze, Germany) and 1 microliter 250 microgram per ml trypsin (Worthington, Lakewood, NJ, USA) were added. The supernatants were digested for 3 hours at 37 degrees Celsius. The digested supernatants were mixed in a 1:1 ratio with 10 milligram per ml recrystallized alpha-cyano-4-hydroxycinnamic acid (?CHCA) (Laser Bio Labs, Sophia-Antipolis Cedex, France) solution; from these mixtures 1 microliter was spotted on the MALDI target. The spotted samples were analyzed with MALDI MS and MALDI MS/MS on a vMALDI LTQ mass spectrometer (Thermo electron, Orlando, FL, USA).

The obtained MS/MS spectra, converted to peptide sequences, were compared to available protein sequences in pre-constructed FASTA protein databases, to identify individual enzymes.

RNA isolation

Aspergillus niger mycelium formed on 20ml liquid selective regeneration medium (6 g/l NaNO₃, 0.52 g/l KCl, 1.52 g/l KH₂PO₄, 0.25 g/l KOH, 0.52 g/l MgSO₄·7H₂O, 22 mg/l ZnSO₄·7H₂O, 11 mg/l H₃BO₃, 5 mg/l FeSO₄·7H₂O, 1.7 mg/l CoCl₂·6H₂O, 1.6 mg/l CuSO₄·5H₂O, 5 mg/l MnCl₂·2H₂O, 1.5 mg/l Na₂MoO₄·2H₂O, 50 mg/l EDTA, 341 g/l sucrose, 10 ml Penicillin (5000 IU/ml) Streptomycin (5000 UG/ml) solution (Gibco)) in a petridish was harvested, washed with demineralized water and squeezed between paper towels to remove excessive water. Mycelium was frozen immediately in liquid nitrogen and grinded to a fine powder using mortar and pestle. The resulting powder was transferred to a sterile 50 ml tube and weighed upon which for every 1-1.2 g of ground mycelium 10 ml TRIzol reagent (Invitrogen) was added (max. 25 ml per tube). The mycelial powder was immediately solubilized by vigorous mixing (vortexing, 1 min.), followed by 5 min room temperature incubation with occasional mixing. 0.2 (original TRIzol) volumes of chloroform (thus 2 ml for every 10 ml TRIzol used originally) was added, vortexed and left at room temperature for 10 min. Subsequently, the mixture was centrifuged at 4 degrees Celsius, 6000 g for 30 minutes. The top aqueous phase was transferred to a fresh tube and total RNA was precipitated by addition of 0.5 (original TRIzol) volumes of 100% isopropyl alcohol (thus 5 ml

of isopropyl alcohol for every 10 ml TRIzol used originally). After 10 minutes precipitation at room temperature, the RNA was recovered by centrifugation for 30 minutes at 6000 g. Upon removal of supernatant, the RNA pellet was rinsed with one (original TRIzol) volume of 70% ethanol. After removal of the ethanol, the RNA pellet was air dried. The dried RNA pellet was dissolved in 3 ml GTS (100 mM Tris-Cl, pH 7.5, 4 M guanidium thiocyanate, 0.5 % sodium lauryl sarcosinate) buffer. Additional clean-up of the RNA was performed using Rneasy Maxi Kit (Qiagen) according to suppliers manual. 10 microliter of RNA solution was used to determine quality and concentration of nucleic acids.

Aspergillus niger strains

A. niger WT-1: This *A. niger* strain is CBS513.88 comprising deletions of the genes encoding glucoamylase (*glaA*), fungal amylase and acid amylase. *A. niger* WT 1 was constructed by using the "MARKER-GENE FREE" approach as described in EP 0 635 574 B1. In this patent it is extensively described how to delete *glaA* specific DNA sequences in the genome of CBS 513.88. The first procedure resulted in a MARKER-GENE FREE ? *glaA* recombinant *A. niger* CBS 513.88 strain, possessing finally no foreign DNA sequences at all. Subsequently, the genes encoding fungal amylase and acid amylase were deleted from the genome of *A. niger* WT-1 by the method described in EP 0 635 574 B1.

Vectors used for cloning of the desired constructs as depicted in figures 7 to 9 and 11 to 14 contain backbone features previously described in WO 99/32617.

A. niger WT-GFP: This *A. niger* strain was obtained by transformation of *A. niger* WT-1 with the expression vector pGBFINGFP-2 depicted in figure 7 containing the gene coding for the well-known Green Fluorescent Protein (GFP) (Chalfie, M et al., Science (1994) 263(5148): 802-805) driven by the glucoamylase promoter. After purification on selective medium, a single copy integrant was selected based on PCR analysis.

A. niger WT-PHY: This *A. niger* strain was obtained by transformation of the *A. niger* WT-1 with the expression vector pGBFIN-32 depicted in figure 8 containing the phytase gene PHY (identical to FytA previously described in WO 99/32617) driven by the glucoamylase promoter. After purification on selective medium, a single copy integrant was selected based on PCR analysis.

A. niger WT-PHY-2: This *A. niger* strain was obtained by transformation of the *A. niger* WT-1 with the expression vector pGBFIN-32 depicted in figure 8 containing the phytase gene PHY (identical to FytA previously described in WO 99/32617) driven by the glucoamylase promoter. After purification on selective medium, a multi copy integrant was selected based on PCR analysis.

A.niger WT-vector: This *A. niger* strain was obtained by transformation of the *A. niger* WT-1 strain with empty control vector pGBFIN-40, depicted in figure 9. The empty control vector pGBFIN-40 was constructed by XhoI digestion of pGBFIN-32 depicted in figure 8, and re-ligation of the largest fragment, thereby removing the glucoamylase promoter and the 5' part of phytase gene PHY. A single copy integrant was selected by PCR analysis.

Example 1

Construction of Bacillus reporter plasmids and reporter strains.

1.1 Construction of reporter plasmids pPhtrA-gfp-amyE and pPhtrB-gfp-amyE

First, reporter plasmid pPhtrA-gfp-amyE was constructed as follows. The gene encoding the optimized version of GFP, *gfpmut-1* (P.Cormack, R.H.Valdivia, and S.Falkow, 1996; FACS-optimized mutants of the green fluorescent protein (GFP), Gene 173, 33.), was amplified by PCR, using pSG1151 (P.J.Lewis and A.L.Marston, 1999; GFP vectors for controlled expression and dual labelling of protein fusions in *Bacillus subtilis*, Gene 227, 101.) as template DNA, with the following primers: RN-lacZ-rv (SEQ ID NO: 3), GTGAGCGGATAACAATTTACACAGG; and gfp-terminator-rv,(SEQ ID NO: 4) GTGGCTCAGCTTTTTTAAGGAAGGGAGGCTCTCACCTCCCTTCCCTTTATTTGTAGAGC TCATCCATGCC, resulting in KpnI and Bpu1101I (in bold) sites up- and downstream *gfpmut-1*, respectively. The 913 bp PCR product was ligated in pDL (containing homologous flanking regions of the *amyE* locus for site specific recombination, (G.Yuan and S.L.Wong, 1995; Regulation of groE expression in *Bacillus subtilis*: the involvement of the sigma A-like promoter and the roles of the inverted repeat sequence (CIRCE), J. Bacteriol. 177, 5427), using the KpnI and Bpu1101I sites, resulting in pGFP-amyE. The promoter of *htrA* (*P_{htrA}*)(SEQ ID NO: 1) was amplified by PCR using chromosomal DNA of *B. subtilis* 168 (*Bacillus* Genetic Stock Center ID:1A1) as template DNA. Primers used were: PhtrA-fw,(SEQ ID NO: 5) CGTGAGGTACCGGCTTCTGTTTCTGCC; and Phtr-rv,(SEQ ID NO: 6) CATCACGAAGCTTATCCATCATGTTCACTCCG, introducing KpnI and HindIII sites (in bold), respectively. After digestion with KpnI and Hind3, the 531bp PCR product was ligated into pGFP-amyE, upstream the *gfp* gene resulting in reporter plasmid pPhtrA-gfp-amyE. A map of this plasmid is depicted in figure 1. Plasmid pPhtrB-gfp-AmyE was constructed in a similar way as described above. The primers used for the amplification of the promoter of *htrB* (*P_{htrB}*) (SEQ ID NO 2) were PhtrB-fw (SEQ ID NO 7) GACCGGTACCTCAGGATCTTTCGCC and PhtrA-rv (SEQ ID NO 8) GGCCATCAAGCTTATAATCCATGTTCTTACACTCC.

1.2 Construction of the *Bacillus* reporter strains VT210A and VT210B

The resulting plasmids pPhtrA-gfp-amyE and pPhtrB-gfp-amyE were introduced into *B. subtilis* DB104, a *DaprE*, *DnprE* double mutant (F.Kawamura and R.H.Doi, 1984; Construction of a *Bacillus subtilis* double mutant deficient in extra cellular alkaline and neutral proteases, J. Bacteriol. 160, 442). Chloramphenicol resistant transformants were checked for site-specific integration by double crossover at the *amyE* locus by PCR on chromosomal DNA. To reduce negative feedback regulation by HtrA of *P_{htrA}*, (D.Noone, A.Howell, R.Collery, and K.M.Devine, 2001; YkdA and YvtA, HtrA-like serine proteases in *Bacillus subtilis*, engage in negative auto regulation and reciprocal cross-regulation of ykdA and yvtA gene expression, J. Bacteriol. 183, 654) the new strain was transformed with chromosomal DNA of *B. subtilis* BV2003, which has a disrupted *htrA* gene by integration of pMutin2 (H.L.Hyyrylainen, A.Bolhuis, E.Darmon, L.Muukkonen, P.Koski, M.Vitikainen, M.Sarvas, Z.Pragai, S.Bron, J.M.van Dijl, and V.P.Kontinen, 2001; A novel two-component regulatory system in *Bacillus subtilis* for the survival of severe secretion stress, Mol. Microbiol. 41, 1159). Transformants were selected for erythromycin resistance and checked by PCR on chromosomal DNA for *htrA* disruption. The obtained strains were designated VT200A and VT200B.

In order to increase transformation efficiency of the VT200A strain, a *BsuMR* disruption construct was designed. The *BsuMR* system is formed by an operon consisting of three genes, *ydiR*, *ydiS* and *ydjA*, of which each is essential for *BsuMR* functioning (P.J.Lewis and A.L.Marston, 1999; GFP vectors for controlled expression and dual labeling of protein fusions in *Bacillus subtilis*, Gene 227,101). Flanking region 1 (flr1), a 756 bp region of *ydiR* was amplified using the following primers: flrBsu1F, (SEQ ID NO 9) GAAAGATTGTTTCAGAAGCC; and flr1BsuR, (SEQ ID NO 10) **ACAGCGTTGGGATCCAAGCCC**TTCCATTTTGGACATTTGG and chromosomal DNA of *B. subtilis* as a template. The spectinomycin resistance marker was amplified using pDG1726 (A.M.Guerout-Fleury, K.Shazand, N.Frandsen, and P.Stragier, 1995; Antibiotic-resistance cassettes for *Bacillus subtilis*, Gene 167, 335) as template DNA and with the following primers: specBsu-F, (SEQ ID NO 11) **GGGCTTGGATCCCAACGCTGT**CGACGTTGTAAAACGACGG; and specBsu-R, (SEQ ID NO 12) CGCATAGCTTTCCGGTCCGCGCAGCTATGACCATGATTACGC. The 1319 bp PCR product and the flrBsu1 fragment were used in a second PCR in which both fragments served as one template upon recombination of their homologous ends introduced by primers flrBsu1-R and specBsu-F (in bold). Using primers flrBsu1-F and specBsu-R, a 2075 bp PCR fragment, flr1-spec^r, was obtained, which was sub cloned in pCR-XL-TOPO

(Invitrogen) resulting in pCT-XL-TOPO-flr1-spec. Flanking region 2, a 739 bp region covering 24 bp of the 3' end of *ydiS* (encoding a DNA endonuclease) and a large part of the adjacent *ydjA* (unknown function), was amplified using primer flrBsu2-F, (SEQ ID NO 13) GATGATTTCGTCTTTTGTAGTG; and primer flrBsu2-R, (SEQ ID NO 14) CGATCAACATATGTGTTTCACG and chromosomal DNA of *B. subtilis* as a template. The PCR product was cloned in pCR-XL-TOPO resulting in pCR-XL-TOPO-flr2. Making use of the multiple cloning site of the pCR-XL-TOPO vector, flr1-spec^r was restricted from pCR-XL-TOPO-flr1 using XhoI and NsiI and ligated into pCR-XL-TOPO-flr2, digested with XhoI and PstI (compatible to NsiI). This resulted in the deletion construct pTOPO-flr1-spec^r-flr2, which was introduced into strain VT200A. Spectinomycin resistant transformants were checked for the *ydiS* deletion by PCR on chromosomal DNA. The obtained, final reporter strains were designated VT210A and VT210B and used for further experiments. A general outline of the construction of these strains is depicted in figure 2.

1.3 Construction of expression vectors pUBnpr2 and pUBBAG:

The construction of plasmid pUB110 and its sequence was described before (McKenzie T., Hoshino T., and Sueoka, N. 1986. The nucleotide sequence of pUB110; some salient features in relation to replication and its regulation. *Plasmid* **15**: 93-103). Cloning of the *Bacillus amyloliquefaciens* α-amylase gene *amyQ* (Genbank accession number J 01542) in pUB110 lead to plasmid pKTH10 as described before (Palva, I. 1982. Molecular cloning of the alpha-amylase gene of *Bacillus amyloliquefaciens* and its expression in *B. subtilis*. *Gene* **19**: 81-87).

Cloning of the neutral protease gene *npr* (Genbank accession number K 02497) and β-glucanase gene *bag* (Genbank accession number M15674) both from *Bacillus amyloliquefaciens* into vector pUB110 was performed essentially according to the method used for the cloning of *amyQ*, resulting in plasmids pUBnpr2 and pUBBAG as depicted in figures 3 and 4.

1.4 Transformation of expression constructs pKTH10, pUBnpr2 and pUBBAG into *Bacillus* reporter strain VT210A:

Transformation of *Bacillus subtilis* strain VT210A was essentially performed as described before making use of the natural competence of this organism (Anagnostopoulos, C. and Spizizen, J. 1961. Requirements for transformation in *Bacillus subtilis*. *J. Bacteriology* **81**: 741-746). After DNA incubation the cells were spread on 2 x TY agar containing 0.1% starch and the antibiotic kanamycin (20 µg/ml) as a selective agent for transformation. Plates were incubated over night at 37° C. When plasmid pKTH10 was

involved in transformation the colonies were covered with a lugol solution to detect the amylase secreting cells.

Example 2

Analysis of secretion stressed cells by fluorescent activated cell assays

2.1 Selection of secretion stressed cells by fluorescence activated cell sorting:

The plasmids pUB110 (empty control vector) and pKTH10 (encoding *Bacillus amyloliquefaciens* α -amylase gene *amyQ*) were used in a ratio of 1:20 for co-transformation of *Bacillus subtilis* strain VT210A. One part of the cells was spread on selective TY agar plates, and to the other part 2 x TY broth containing kanamycin, was added and the cells were cultivated over night. The separate plasmids were transformed as well, performing the same operations as for the co-transformation. The next day, cultures were analyzed for GFP fluorescence on a flow cytometer type FACS (Mo-Flo of Dako-cytomation) by a skilled operator. A blue laser of 80 mWatt, and 488nm was used. A typical FACS experiment comprised analysis of 20.000 events (cells). First, fluorescence signals of cells of VT210A, pUB110 (no secretion stress) and of VT210A, pKTH10 (secretion stress due to overproduction of the α -amylase AmyQ) were analyzed separately (figures 5A and 5B) and as a mix (figure 5C). On basis of these results a cell-sorting limit was manually set (dashed line in figure 5D), such that cells demonstrating GFP fluorescence higher than the cell-sorting limit, were sorted. The co-transformation culture was applied to the flow cytometer and sorted cells were collected in a tube. To verify that the sorted fluorescent population had been enriched for α -amylase secreting cells, different dilutions of the collected cells were spread on selective agar containing starch. The same was done with the original overnight input culture, which was not sorted by FACS. After growth, the colonies were covered by a lugol solution to visualize halo formation caused by α -amylase activity. As shown in table 1, about 90% of the sorted cells indeed secreted α -amylase, whereas in the input culture only 0.5 – 0.7% of the cells secreted α -amylase. This result clearly demonstrates the powerful selection tool of the present invention by isolation of protein secreting cells using discrimination in fluorescent signal between normal cells and cells stressed by secretion.

Used cells	Number of cfu	Number of amy ⁺	Dilution	% amy ⁺
Pre-FACS of over night grown co-transformation	>2000	14	10 ⁻⁴	< 0.7
	600	3	10 ⁻⁵	0.5

Sorted cells of over night grown co-transformation	111	104	10^{-1}	93.7
	9	8	10^{-2}	88.9

Table 1: enrichment for amy⁺ colony forming units after flow cytometry.

2.2 Selection of secretion stressed cells by fluorescent cell scanning:

The plasmids pUB110 (empty control vector) and pKTH10 (encoding *Bacillus amyloliquefaciens* α -amylase gene *amyQ*) were transformed separately to *Bacillus subtilis* strain VT210B. One part of the cells was spread on selective TY agar plates, and to the other part 2 x TY broth containing kanamycin, was added and the cells were cultivated over night. Cultures were analysed by a skilled operator on a flow cytometer (Coulter Epics XL-MCL, Beckman Coulter). A typical experiment comprised analysis of 20.000 events (cells). Fluorescent signal was detected using a FITC-filter and setting of the photo multiplier voltage was between 700 and 800 Volts. The difference in fluorescent signal between the normal cells and cells stressed by secretion is depicted in figure 6.

2.3 Selection for secretion stressed VT210A cells using a fluorimeter:

Transformed reporter strain cells were inoculated in 2 xTY broth in the presence of kanamycin and grown over night at 37° C. The DNA's used for transformation were equal amounts of the plasmids pUB110, pKTH10, pUBnpr2 and pUBBAG. After growth to stationary phase the cultured cells of the four transformations were divided 8-fold in equal volumes in a 96 well micro titer plate format. The wells were analyzed for fluorescence in a fluorimeter (Molecular Devices, type Spectra MAX Gemini). Excitation is at 490 nm, emission at 510 nm, and cutoff at 495 nm. A standard t-Test assuming unequal variances was performed. Table 2 demonstrates that non-stressed cells (pUB110) were distinguished from the stressed cells (pKTH10, pUBnpr2 and pUBBAG) by a clear significant (P two-tail < 0.05) difference in fluorescent signal.

	pUB110 cells	Ratio
pKTH10 cells	P = 1,05899E-13	7,84
pUBnpr2 cells	P = 4,63415E-11	1,58
pUBBAG cells	P = 7,17216E-14	1,77

Table 2 t-Test assuming unequal variances. P-values are given for empty host cells (pUB110) versus secretion stressed (pKTH10, pUBnpr2, pUBBAG) cells.

The ratio expresses the mean (n=8) fluorescent value of stressed cells over the mean (n=8) fluorescent value of empty host cells.

Example 3

Genome-wide expression analyses of *A. niger* to identify secretion-induced promoters

5 Genome-wide expression analyses were performed to identify promoter sequences that are highly expressed in *A. niger* WT-PHY, and show no or low expression in *A. niger* WT-GFP, *A. niger* WT-vector and *A. niger* WT-1. To simulate future application experiments protoplasts were obtained of *A. niger* WT-1, *A. niger* WT-GFP, *A. niger* WT-PHY and *A. niger* WT-vector strains. An amount of 20 ml liquid selective regeneration medium supplemented with 50 mg/l phleomycin (Invitrogen) in case of *A. niger* WT-GFP, *A. niger* WT-PHY and *A. niger* WT-vector, was inoculated in a petri dish with 100 microliter 10E7/ml protoplasts. After a growth period of 3 days at 30 degrees Celsius (static), a mycelial pellet had formed at the surface of the liquid medium which was used for total RNA isolation. cDNA synthesis, labelling of the cDNA and hybridisation on Affymetrix *A. niger* GeneChips™
10 was performed according to suppliers protocol. Subsequent expression analysis resulted in identification of 5 genes that had an expression level in the *A. niger* WT-PHY strain of at least 8.5 times the basal expression level in *A. niger* WT-1, *A. niger* WT-GFP and *A. niger* WT-vector. To confirm the findings, Northern blot analysis was performed.

Example 4

Northern analysis

20 To corroborate the results of the Affymetrix *A. niger* GeneChips™ experiments, Northern blot analysis was performed. For this, total RNA was isolated from *A. niger* strains WT-1, WT-GFP, WT-FYT and WT-vector. RNA was denatured, separated by gel electrophoresis on a 1% agarose gel and transferred onto a nylon membrane (Hybond-N+, Amersham Biosciences) by capillary blotting according to the manufacturer's instruction.
25 The RNA was UV cross-linked to the nylon membranes on a GS Gene Linker™ (BioRad, C3 setting, 150 mJoule). DNA for the generation of ³²P-labeled DNA probes was isolated by PCR using genomic DNA of *A. niger* WT-1 as a template and using the primers listed in Table 3 (SEQ ID NO's 50 to 59). Northern blots were hybridised with randomly primed ³²P-labeled DNA probes, representing the secretion-induced genes (SEQ ID NO's 20 to 24), according to suppliers instructions (RadPrime DNA Labeling System, Invitrogen). The results of the Northern analyses are shown in figure 10. The hybridisation pattern of the Northern blots confirm the data obtained by expression analyses of the Affymetrix *A. niger* GeneChips™ data; low expression of the secretion-induced genes in *A. niger* strains WT-1, WT-GFP, WT-vector and high expression in *A. niger* WT-PHY.
30
35

Primer	Primer sequence	Gene sequence
SEQ ID NO: 50	CAACACCTACGGCGTCAAGG	SEQ ID NO:20
SEQ ID NO: 51	TAACCCGCGGGAGATGCTGTT	SEQ ID NO:20
SEQ ID NO: 52	CGATCCGGAATTCCTCCTG	SEQ ID NO:21
SEQ ID NO: 53	GCCGCCGACTCAGCAGTAT	SEQ ID NO:21
SEQ ID NO: 54	GTGCATGAACAACACGGAGA	SEQ ID NO:22
SEQ ID NO: 55	ACTCAGTTGTGCGCCTCTCG	SEQ ID NO:22
SEQ ID NO: 56	GTCGAACTCGGCGTTTCTCC	SEQ ID NO:23
SEQ ID NO: 57	CCGGTGAGGAGGTGCTTTTC	SEQ ID NO:23
SEQ ID NO: 58	AGCATCTCGCCATCCATCAG	SEQ ID NO:24
SEQ ID NO: 59	GCCTTCGGACCATGGTATCC	SEQ ID NO:24

Table 3: Primers used to isolate DNA for ^{32}P -labeled probes.

Example 5

Cloning of GFP-BLE secretion-induced reporter constructs.

A fusion of GFP (Green Fluorescent Protein (Chalfie, M et al., Science (1994) 263(5148): 802-805) and phleomycin binding protein (BLE, present in pGBFIN-14 previously described in WO99/32617) was constructed by means of PCR. The GFP fragment was amplified by PCR using the primer set consisting of SEQ ID NO: 62 and SEQ ID NO: 63. The BLE fragment was amplified by PCR using the primer set consisting of SEQ ID NO: 64 and SEQ ID NO: 65 with pGBFIN-14 as template. The overlapping PCR fragments obtained were used in a fusion PCR with primers SEQ ID NO: 62 and SEQ ID NO: 65, resulting in PCR fusion product GFP-BLE (SEQ ID NO: 61). The GFP-BLE fusion was digested with restriction enzymes *PacI* and *Ascl* and ligated in the *PacI*, *Ascl* linearized cloning vector pGBFIN-2 (described in WO99/32617 and EP98/08577) thereby placing the GFP-BLE fusion under control of the glucoamylase promoter, the resulting vector pGBFINBLE-2 is depicted in figure 11.

Promoter sequences of the secretion-induced genes were isolated by PCR using genomic DNA from *A.niger* WT-1 as a template and primers as displayed in Table 4.

Primer	Primer sequence	Promoter	PCR product
SEQ ID NO:40	ACTTCATTAATTAAGTTGATTGAGGTAGAGATGAGTTTG	P1, SEQ ID NO:20	SEQ ID NO:35
SEQ ID NO:41	ACTTCACTCGAGATAATGTAGGCGACAAAGTAGCC	P1, SEQ ID NO:20	SEQ ID NO:35
SEQ ID NO:42	ACTTCATTAATTAATTTGACACTTAACGTGGTAAGG	P2, SEQ ID NO:21	SEQ ID NO:36

SEQ ID NO:43	ACTTCACTCGAGTTCTCACAAGCAAATCCGAGA	P2, SEQ ID NO:21	SEQ ID NO:36
SEQ ID NO:44	ATCCTACTCGAGTGGTGCCTGTGAACGAGGTCA	P3, SEQ ID NO:22	SEQ ID NO:37
SEQ ID NO:45	ACTTCATTAATTAATGTGGATTTGGATAGTAATTAAG	P3, SEQ ID NO:22	SEQ ID NO:37
SEQ ID NO:46	ATCCTATTAATTAAGATGGTTTGGTGATAACAGAACA	P4, SEQ ID NO:23	SEQ ID NO:38
SEQ ID NO:47	ACTTCACTCGAGGTCCACACTCTAATTGGGATGA	P4, SEQ ID NO:23	SEQ ID NO:38
SEQ ID NO:48	ATCCTATTAATTAAGTTGCATCTGCGACAGACAGT	P5, SEQ ID NO:24	SEQ ID NO:39
SEQ ID NO:49	ACTTCACTCGAGGCCTTTGGAGAATGTAATATCCC	P5, SEQ ID NO:24	SEQ ID NO:39

Table 4: Overview of primers used for amplification of corresponding promoters

The resulting PCR fragments (SEQ ID NO's: 35 to 39) were digested with restriction enzymes XhoI and PaeI and ligated in the XhoI, PaeI linearized cloning vector pGBFINBLE-2 depicted in figure 11, thereby removing PgpA-amdS and replacing the glucoamylase promoter with the respective secretion-inducible promoter, Pind, resulting in pGBFINGFPBLE-1 to 5. pGBFINGFPBLE-1 is depicted in figure 12, and is representative for pGBFINGFPBLE-2 to 5. The fusion constructs of the respective inducible promoters with GFP-BLE was isolated by XhoI and FseI digestion of pGBFINGFPBLE-1 and subsequently ligated in the XhoI, FseI linearized *A. niger* expression vector pGBTOPSEL-1 depicted in figure 13. The resulting secretion reporter constructs, pGBTOPGFPBLE-1 to 5, contain a fusion of GFP-BLE under the control of respectively secretion-inducible promoter P1, P2, P3, P4, P5. In figure 14 pGBTOPGFPBLE-1 is depicted as representative for pGBTOPGFPBLE-1 to 5. These vectors were used to transform *A.niger* WT-1 in example 6 and 7.

Example 6

Overexpression of extracellular protein confers phleomycin resistance to *A. niger* expressing a GFP-BLE reporter construct.

Transformation of the secretion reporter constructs pGBTOPGFPBLE-1 to 5 to *A. niger* WT-PHY-2 overexpressing the secreted protein phytase was performed according to the procedure described in Materials and Methods. Transformants were selected on selective regeneration plates supplemented with various amounts of phleomycin (0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 microgram per ml) as selective agent. All 5 reporter constructs enabled direct selection of transformants on the phleomycin containing plates whereas the negative control, WT-PHY-2 wild-type, was not able to grow on any of the plates containing phleomycin. The secretion reporter constructs pGBTOPGFPBLE- 1 to 5 enabled growth on respectively 50, 35, 50, 35, 25 microgram per ml phleomycin. Thereby demonstrating that

the secretion-inducible promoters of these constructs are functional and direct selection on phleomycin after transformation is possible.

Example 7

Selection of *A.niger* clones expressing extracellular proteins after co-transformation of a cDNA library and the GFP-BLE reporter construct.

Aspergillus niger WT-1 was co-transformed using the secretion reporter constructs pGBTOPGFPBLE-1 to 5 in combination with a defined *A. niger* cDNA library encoding for 50 extracellular proteins and 50 intracellular proteins. Co-transformation events were confirmed by colony PCR. Spores of co-transformants containing the secretion reporter construct in combination with an expression library construct were grown on selective plates supplemented with different amounts of phleomycin (0, 5, 25, 50, 100, 175 microgram per ml). An average increase in phleomycin resistance was demonstrated for co-transformants containing the secretion reporter construct in combination with a library construct encoding an extracellular protein compared to transformants with the secretion reporter construct in combination with a library construct encoding an intracellular protein. Graphics representing these results for secretion sensitive promoters P4 and P5 are shown in figure 15 and 16. These results demonstrate that an increase in phleomycin resistance of at least up to 20-fold was achieved.

Diversity of extracellular proteins produced by the co-transformants which contained secretion reporter constructs comprising promoters P1 or P3 in combination with a library construct encoding an extracellular protein was demonstrated. Co-transformants were individually grown in microtiterplates and on day 6 the supernatant of these cultures was subjected to protein analysis by E-PAGE™ High-Throughput Pre-Cast Gel System (Invitrogen) and mass spectrometry (MALDI MS/MS).

The E-PAGE™ 96, 6% Gels (Invitrogen) contain 96 sample lanes in microtiterplate format and 8 marker lanes for protein separation. Supernatant and marker (E-PAGE™ SeeBlue® pre-stained standard (Invitrogen)) were loaded on the gel according to suppliers manual. After electrophoresis of the gel, SimplyBlue™ Safestain (Invitrogen) was used for staining according to suppliers manual. The scan of the E-PAGE™ gel was edited by the E-editor program provided by Invitrogen, which aligns the protein samples in microtiterplate format. Extracellular proteins produced by co-transformants with secretion sensitive promoters P1 and P3 are displayed in figure 17 and 18, and are summarized in Table 5. These results demonstrate that up to 77% of the analysed phleomycin resistant colonies secreted proteins in amounts detectable by E-PAGE. Based on the size of the observed protein fragments on the gels, the diversity of the clones transformed with the defined library

was estimated. These results demonstrate that a diversity of up to 35% could be observed based on E-PAGE analysis.

The mass spectrometry results are depicted in Table 5. Up to 58% of the analysed phleomycin resistant colonies secreted proteins in amounts detectable by mass spectrometry. Diversity was estimated based on the identity of the detected peptide sequences. These results demonstrate that a diversity of up to 100% could be observed using mass spectrometry.

Assay	Promoter	Transformants analysed	Positive		Diversity	
			No	%	No	%
MALDI	1	26	14	54%	14	100%
MALDI	3	31	18	58%	12	67%
E-PAGE	1	26	20	77%	7	35%
E-PAGE	3	31	20	65%	6	30%

Table 5: Summary of E-PAGE and mass spectrometry results. The assay used for analysis is depicted in the first column; the second column identifies the secretion sensitive promoter-reporter construct; the third column depicts the amount of transformants that were analysed by either assay; the fourth and fifth column depict the amount of colonies scored positive for secretion of a protein and the respective positive percentage; the sixth and seventh column depict the number of colonies secreting distinct proteins and the respective diversity percentage.

Applicant's or agent's file reference number 24403WO	International application No.
---	-------------------------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description firstly mentioned on page 14 line 10	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet <input type="checkbox"/>
Name of depositary institution CENTRAAL BUREAU VOOR SCHIMMELCULTURES	
Address of depositary institution (including postal code and country) Uppsalaalan 8 P.O. Box 85167 NL-3508 AD Utrecht The Netherlands	
Date of deposit 10 August 1988	Accession Number CBS 513.88
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
We inform you that the availability of the microorganism identified above, referred to Rule 13bis PCT, shall be effected only by issue of a sample to an expert nominated by the requester until the publication of the mention of grant of the national patent or, where applicable, for twenty years from the date of filing if the application has been refused, with drawn or deemed to be withdrawn.	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only <input type="checkbox"/> This sheet was received with the international application Authorized officer	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer
--	---

CLAIMS

1. A promoter DNA sequence such as:

- 5 (a) a DNA sequence as presented in the following list: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23 or SEQ ID NO:24,
(b) a DNA sequence capable of hybridizing with a DNA sequence of (a),
(c) a DNA sequence being at least 50% homologous to a DNA sequence of (a),
(d) a variant of any of the DNA sequences of (a) to (c), or
10 (e) a subsequence of any of the DNA sequences of (a) to (d).

2. A DNA construct comprising a promoter DNA sequence according to claim 1 operatively associated with a reporter gene conferring a selectable trait, preferably a selection marker gene.

15 3. An expression vector comprising a DNA construct according to claim 2, wherein the DNA construct comprises a DNA fragment, which is homologous to a DNA sequence in a predetermined target locus in the genome of a host cell, preferably the predetermined target locus comprises a highly expressed gene.

20 4. An expression vector comprising a DNA construct according to claim 2, wherein the DNA construct is capable of autonomous maintenance in a host cell, preferably the DNA construct comprises an AMA1-sequence.

25 5. A host cell comprising the DNA construct according to claim 2, or comprising the expression vector according to claim 3 or 4.

30 6. The host cell according to claim 5, further comprising a DNA construct comprising a DNA sequence comprising a coding sequence originating from a DNA library from an organism suspected of being capable of producing one or more proteins of interest.

7. The host cell according to claim 5 or 6, wherein the host cell is a prokaryote or an eukaryote.

8. The host cell according to claim 7, wherein the host cell is selected from the following list: a *Bacillus*, a yeast or a filamentous fungus, preferably an *Aspergillus*, *Penicillium* or *Trichoderma* species.

5 9. The host cell according to claim 8, wherein the *Aspergillus* is an *Aspergillus niger* or *Aspergillus sojae* or *Aspergillus oryzae* species.

10. A method for isolating a DNA sequence comprising a DNA sequence coding for a protein of interest in a host cell, said method comprising the steps of:

- 10 (a) preparing, a first DNA construct comprising a promoter DNA sequence operatively associated with a reporter gene conferring a selectable trait; said promoter DNA sequence being induced when the DNA construct is present in the host cell and when a protein of interest, preferably a secreted protein of interest, is produced by the host cell,
- 15 (b) preparing a second DNA construct comprising a DNA sequence comprising a DNA sequence coding for a protein of interest originating from a DNA library from an organism suspected of being capable of producing one or more proteins of interest,
- (c) transforming a host cell with both DNA constructs prepared in (a) and in (b),
- (d) culturing all the transformed host cells obtained in (c) under conditions conducive to the production of the proteins of interest as present in the DNA library; and
- 20 (e) screening for transformed host cells producing a protein of interest by analysis of the proteins produced in (d).

11. The method according to claim 10, wherein the host cell is first transformed with the DNA construct prepared in step (a) and consecutively is transformed with the DNA construct prepared in step (b).

25

12. The method according to claims 10 or 11, wherein the promoter used in step (a) is the promoter of claim 1.

30

13. A method according to any one of claims 10 to 12, wherein the organism suspected of being capable of producing a protein of interest is a eukaryote or a prokaryote.

14. A method according to any one of claims 10 to 13, wherein the protein of interest is an enzyme.

35

1/18

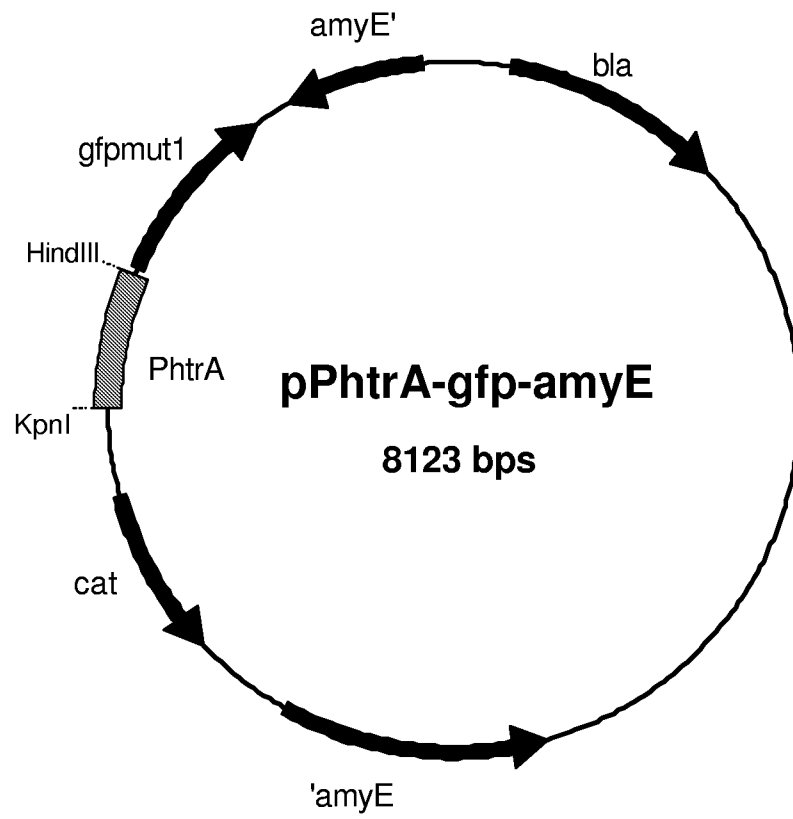


Fig. 1

2/18

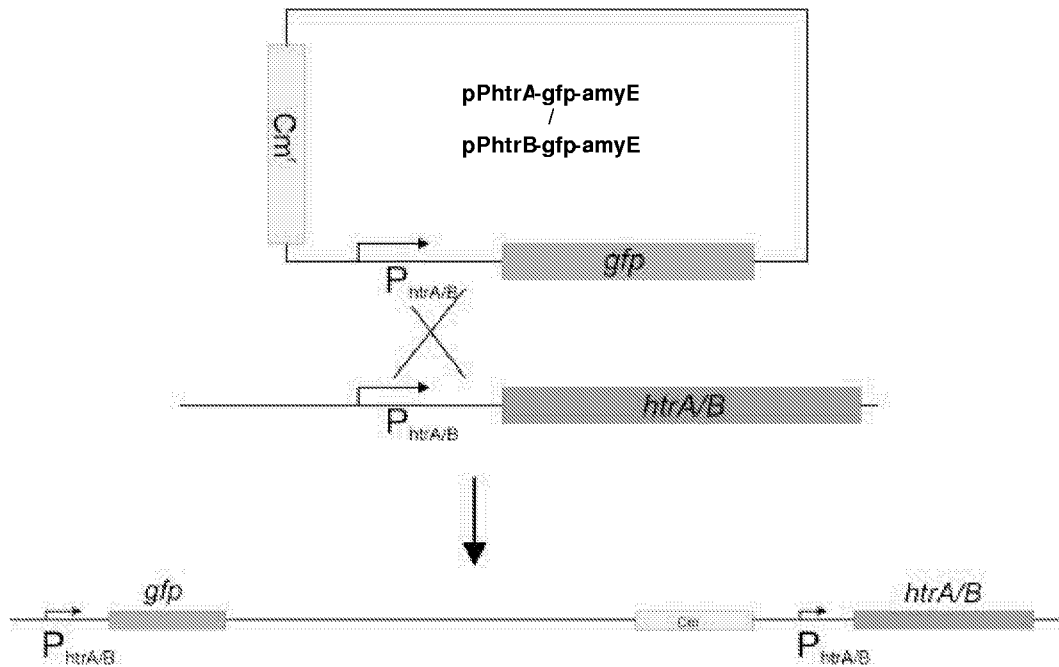


Fig. 2

3/18

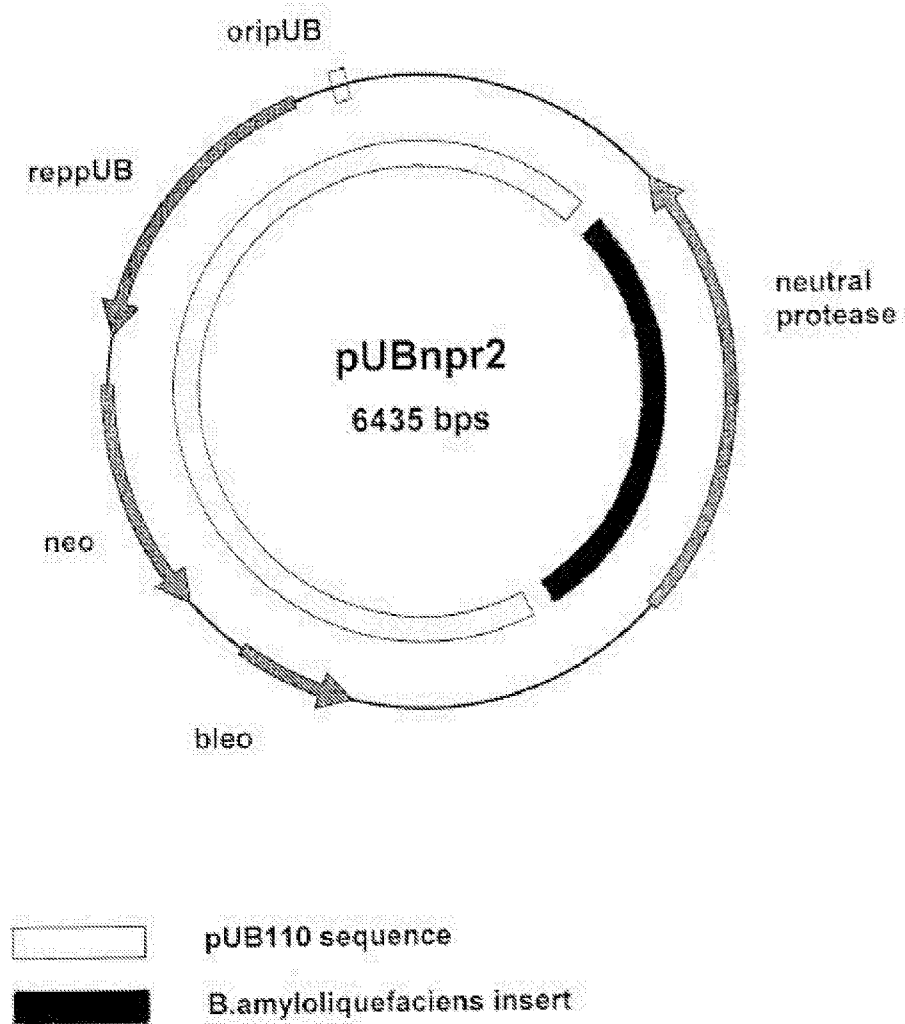


Fig. 3

4/18

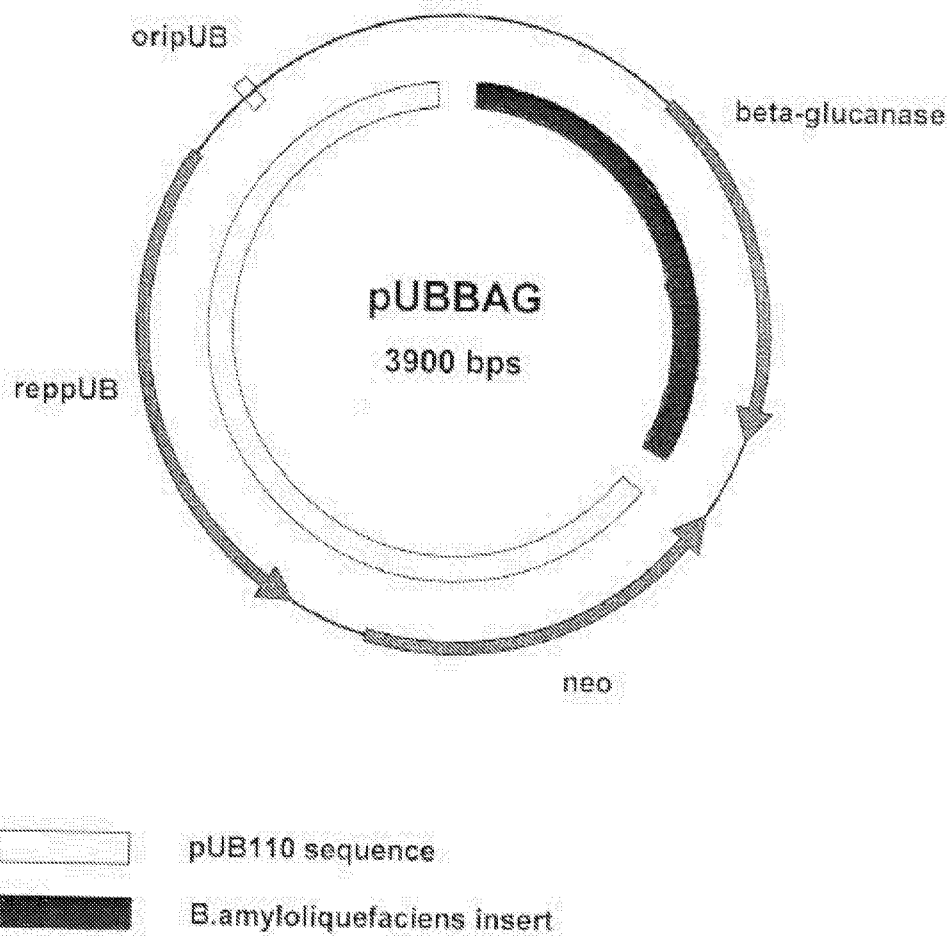
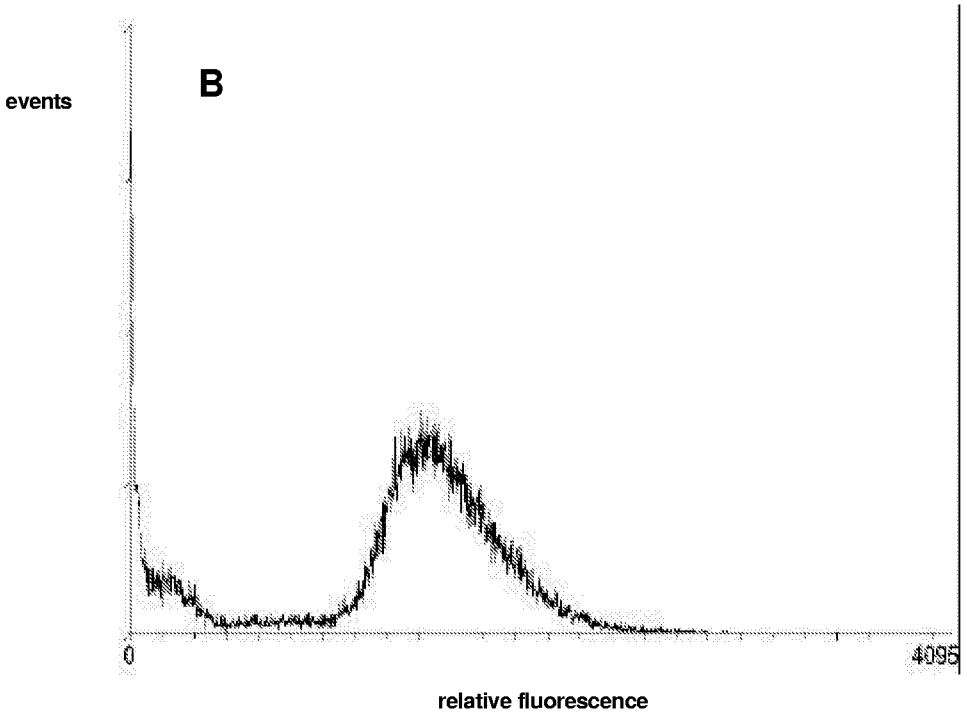
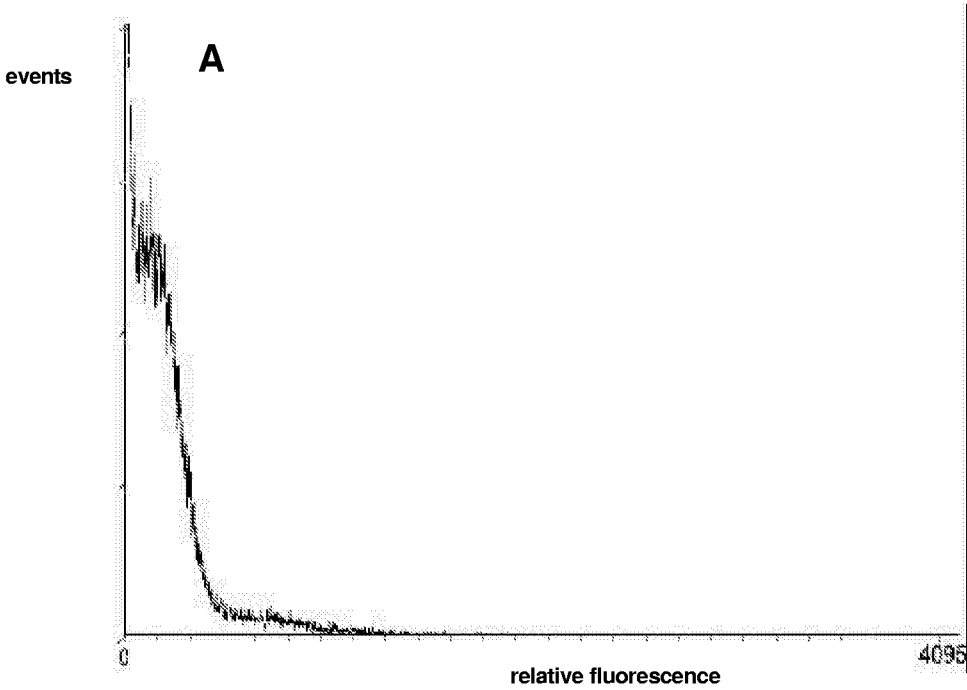


Fig. 4

5/18



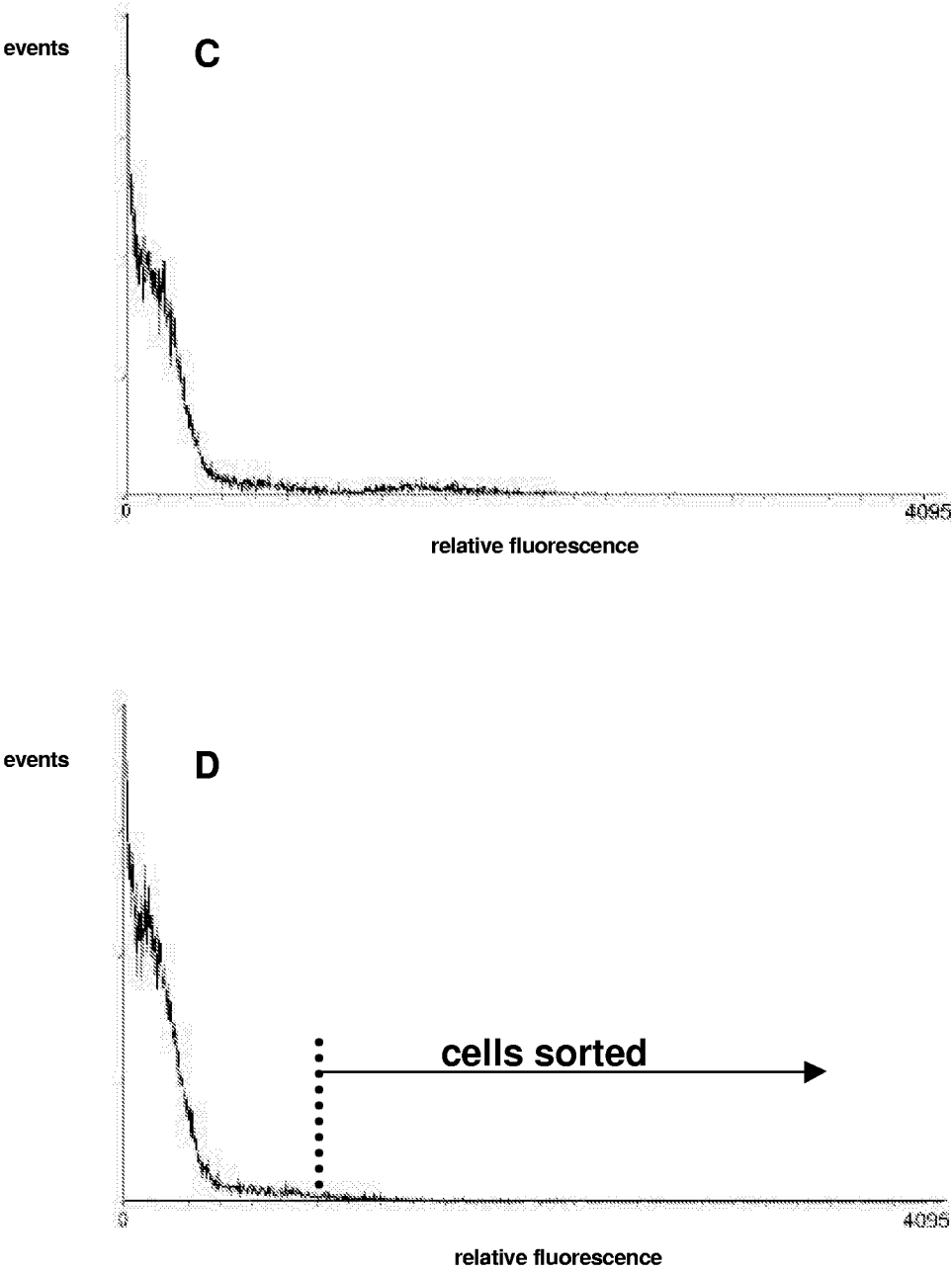


Fig. 5

6/18

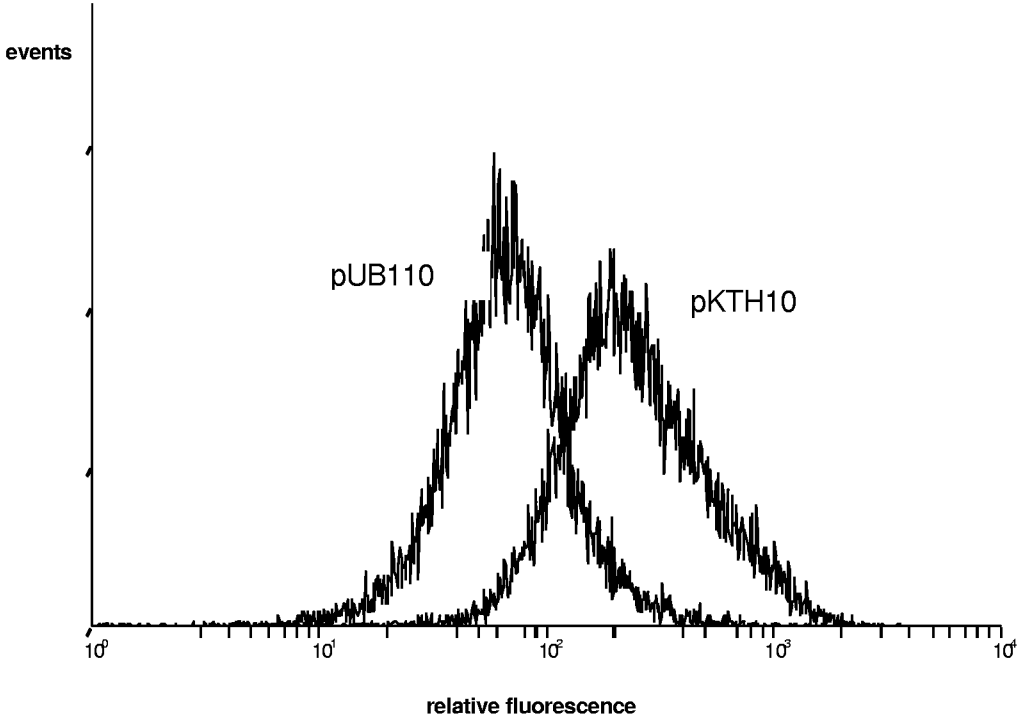


Fig. 6

7/18

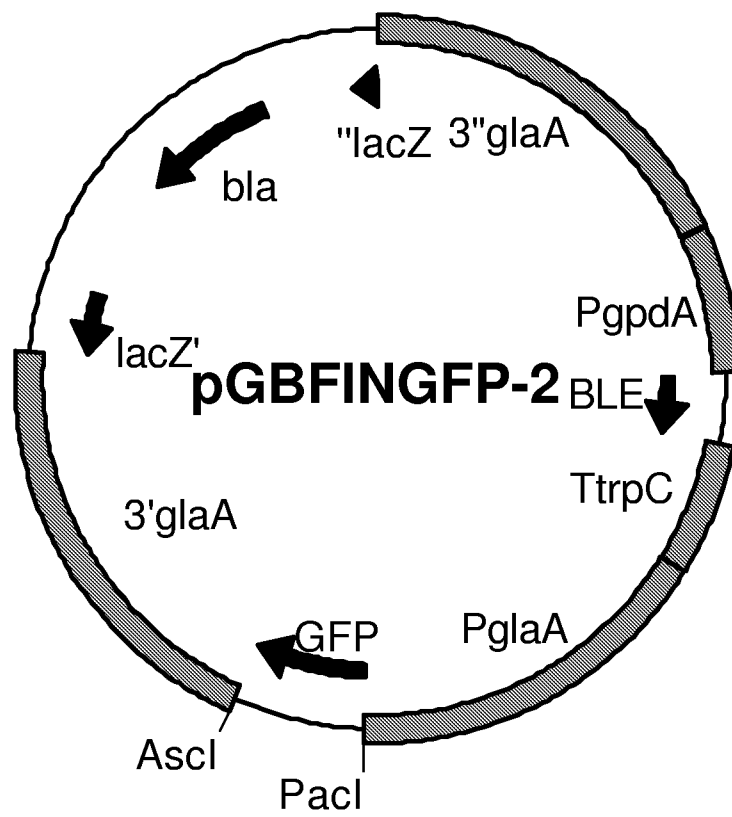


Fig. 7

8/18

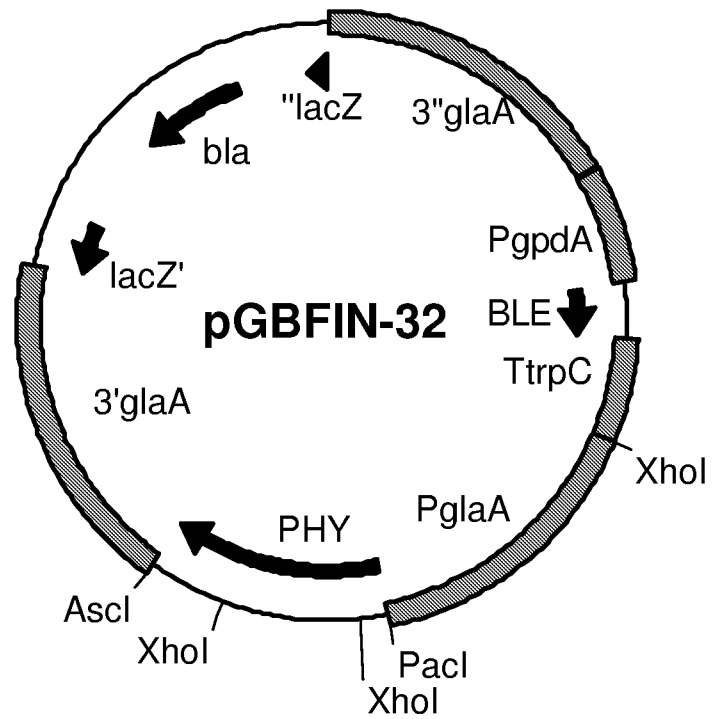


Fig. 8

9/18

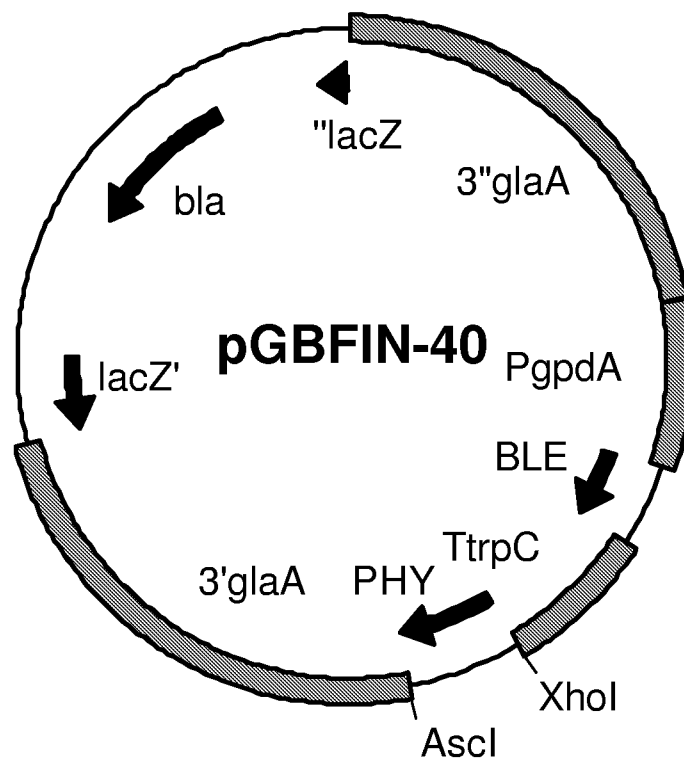


Fig. 9

10/18

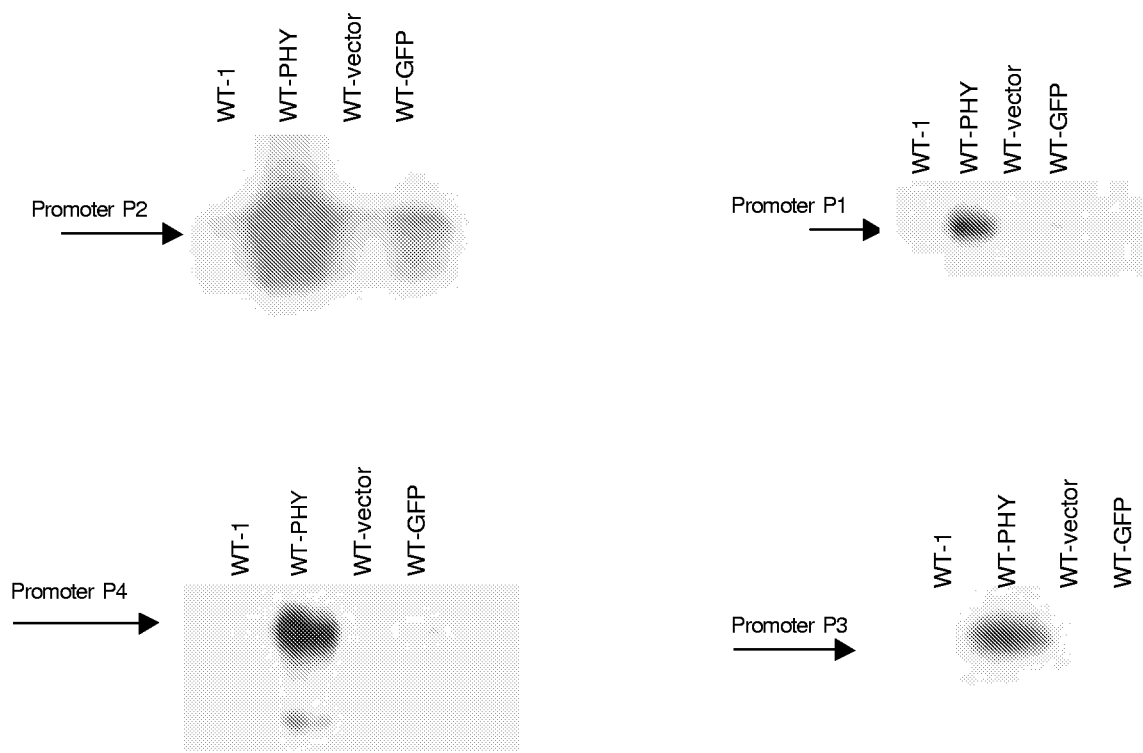


Fig. 10

11/18

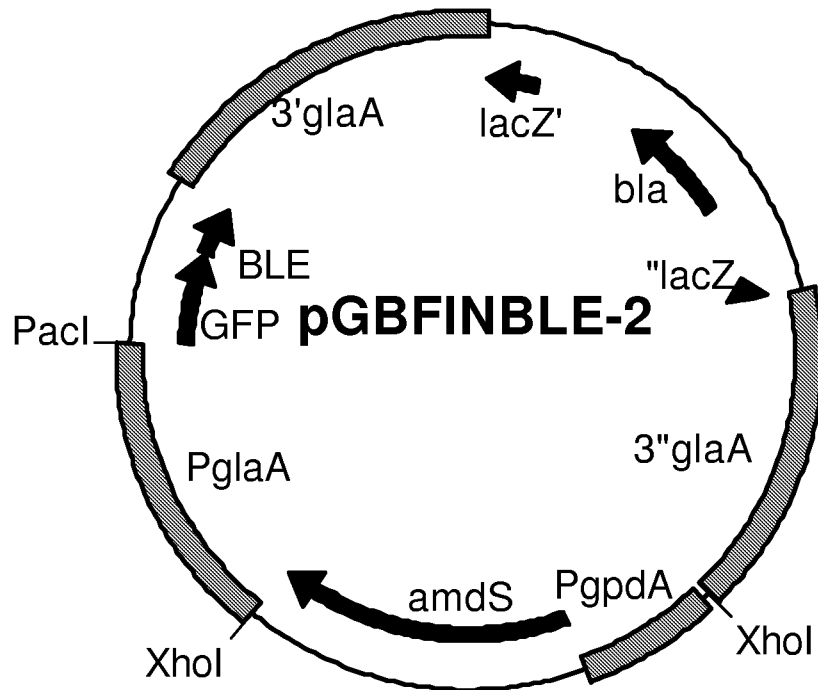


Fig. 11

12/18

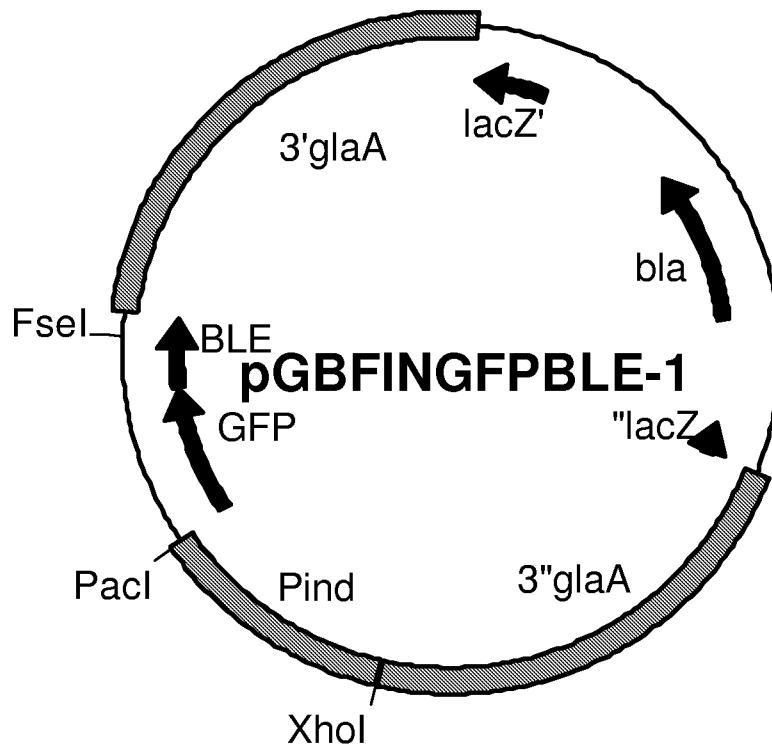


Fig. 12

13/18

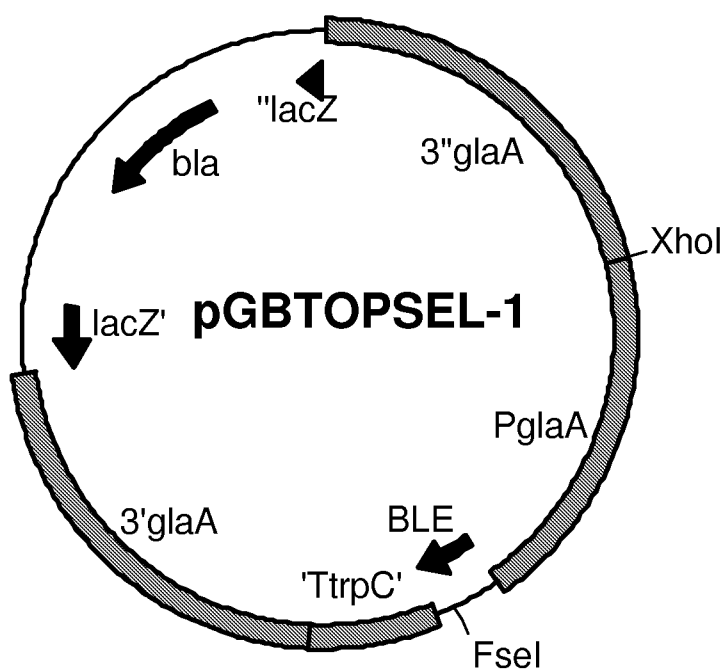


Fig. 13

14/18

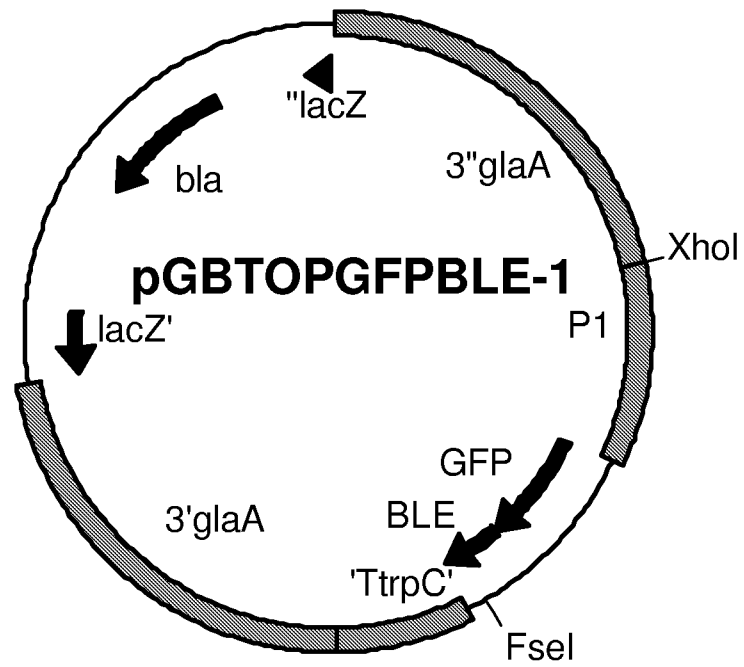


Fig. 14

15/18

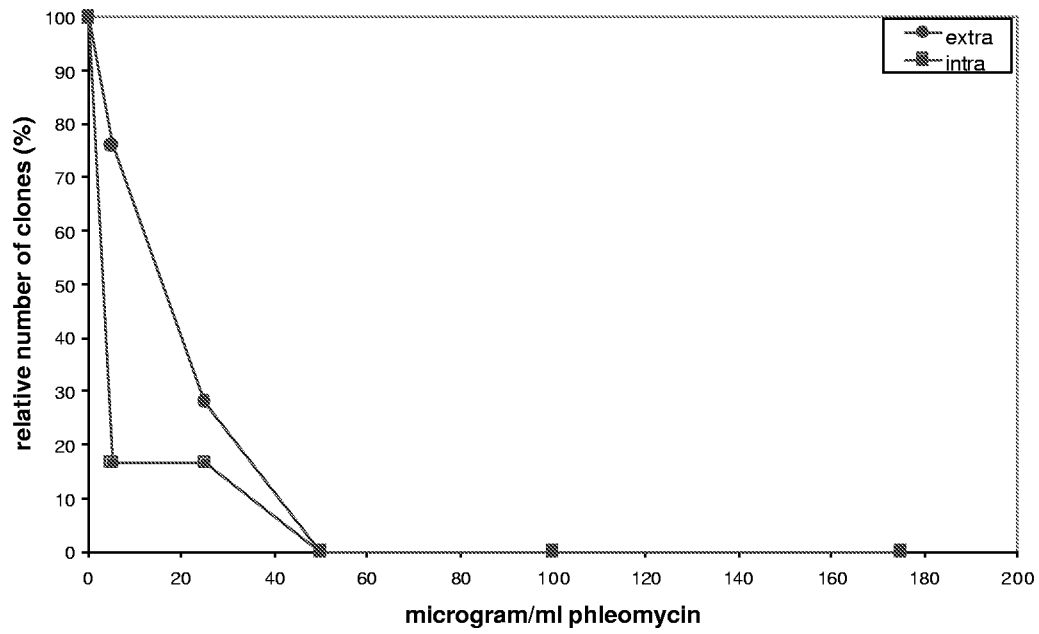


Fig. 15

16/18

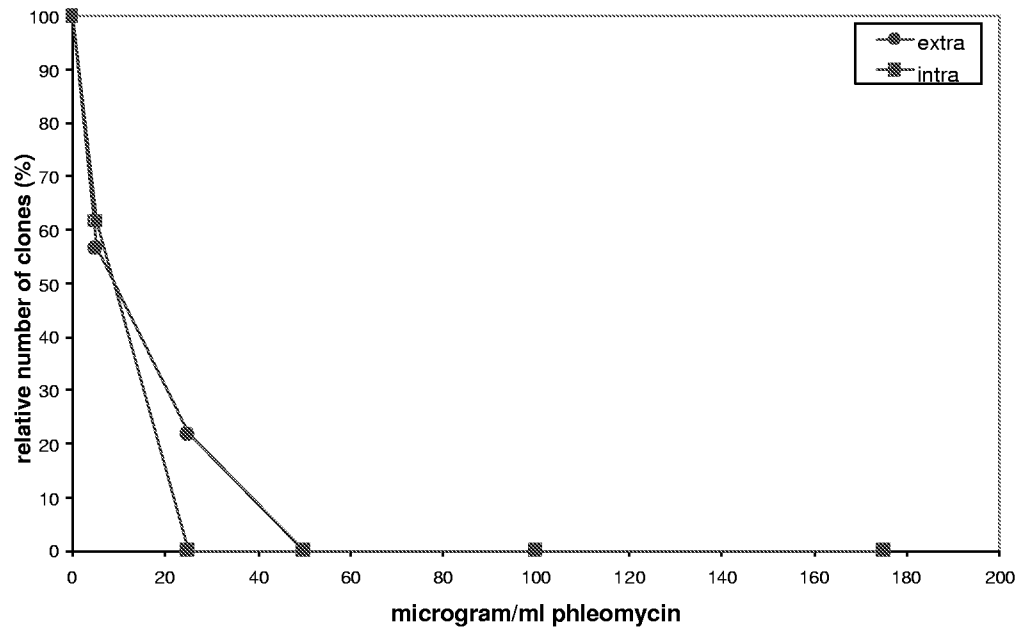


Fig. 16

17/18

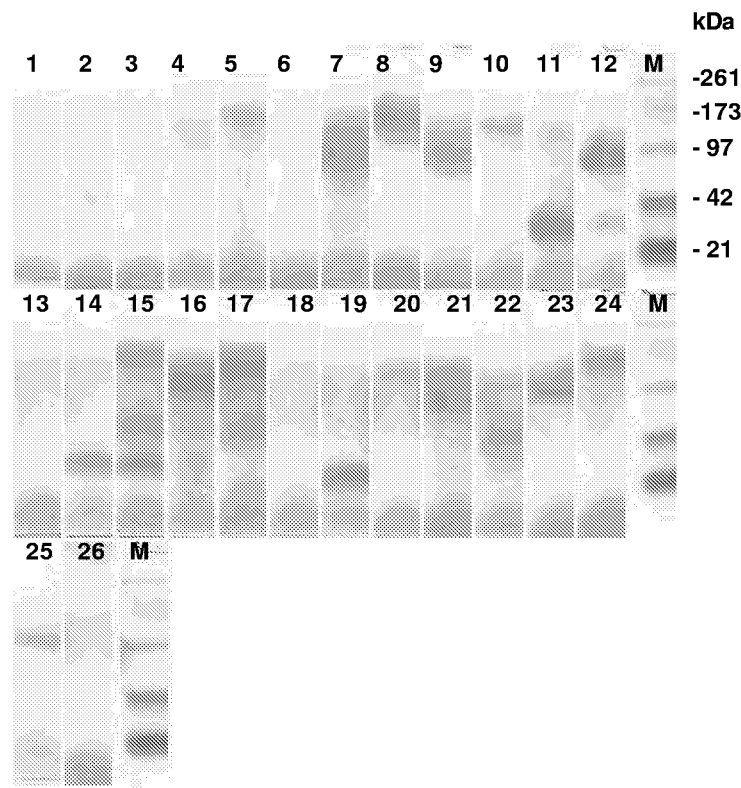


Fig. 17

18/18

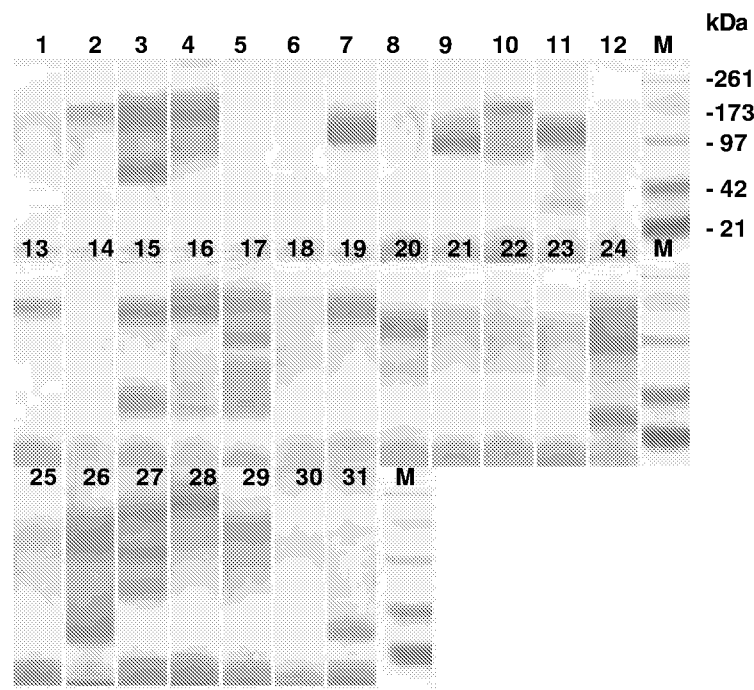


Fig. 18

SEQUENCE LISTING

<110> DSM IP Assets B.V.

<120> IMPROVED METHOD OF EXPRESSION CLONING IN A HOST CELL

<130> 24403WO

<160> 65

<170> XML to WIPO ST.25 Converter - <http://www.biomax.de/>

<210> 1

<211> 168

<212> DNA

<213> Bacillus subtilis

<400> 1

CGACTCAGTC CTTTCATATA CAATATGAAG TGTACCGTTT TCCGCACTTT TTCACAATTT	60
CCCATAATCT TTTTCATTTTT ATCCACAGT TTTTGTATAT GATAAACTCA AGTCATAAAC	120
CTATCAATAT AAATAGACAT GTGAAAATAG AGAAACGGAG TGAACATG	168

<210> 2

<211> 277

<212> DNA

<213> Bacillus subtilis

<400> 2

GGCTCTTCAC ATCCTTTCAA CGTCATTATA AACTAGTTTT AACATACGGC AGGCAATTTT	60
CATAATTTCA CATATTCCTT TCATTTTTAT CCCACAATGT TTGTTTATGA TATGGTTAAG	120
GGAAGAAAGG AAGAGAAAAA GAGCGAGGAA GATGTAGGAT GATCAAACAT ATTCTATTTT	180
GGTGCTTTGC TTTTCTCCTC ATTATTGGGA CAATTGAACT TGTACATGCG ATGAATATGT	240
AACGATAAAT GAATAACCGT TAAAGGAGTG TAAGAAC	277

<210> 3

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 3

GTGAGCGGAT AACAAATTCA CACAGG	26
------------------------------	----

<210> 4

<211> 70

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 4

GTGGCTCAGC TTTTAAAGG AAGGGAGGCT CTCACCTCCC TTCCCTTTAT TTGTAGAGCT 60
CATCCATGCC 70

<210> 5
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 5
CGTGAGGTAC CGGCTTCTGT TTCTGCC 27

<210> 6
<211> 32
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 6
CATCACGAAG CTTATCCATC ATGTTCACTC CG 32

<210> 7
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 7
GACCGGTACC TCAGGATCTT TCGCC 25

<210> 8
<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 8
GGCCATCAAG CTTATAATCC ATGTTCTTAC ACTCC 35

<210> 9
<211> 20
<212> DNA
<213> Artificial sequence

<220>

<223> Primer

<400> 9

GAAAGATTGT TTCAGAAGC

20

<210> 10

<211> 40

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 10

ACAGCGTTGG GATCCAAGCC CTTCCATTTT GGACATTTGG

40

<210> 11

<211> 40

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 11

GGGCTTGGAT CCCAACGCTG TCGACGTTGT AAAACGACGG

40

<210> 12

<211> 42

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 12

CGCATAGCTT TCCGGTCGCC GCAGCTATGA CCATGATTAC GC

42

<210> 13

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 13

GATGATTCGT CTTTTTGTAG TG

22

<210> 14

<211> 21

<212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 14
 CGATCAACAT ATGTGTTTAC G

21

<210> 15
 <211> 1050
 <212> DNA
 <213> *Aspergillus niger*

<400> 15
 ATAATGTAGG CGACAAAGTA GCCGGGCATG CGACCGGAGG ACTCGGTCAG AGGCACGAAT 60
 ATGAGAGGCC AGAACGCCGC ACCCATGTTC CAAGATGTTG TGGCCCAGTA AAGGTTAGGA 120
 AATGGGGTGT TTTGAACGTG GAATCGTTCT TCCATCCAGT TGTTGCCGGA GCCGTAGGAA 180
 CCGGCCGGGA TACCAGTGAG GATGGAAC TAACCTATTAG CACTGCCATC TTCGAGAACA 240
 AGGCTAAGAC ATACATGAAA CAAACGGTGA CCGTGAGAAG CCATTTATAA GAGGTGCTCC 300
 AATTGAACCT GAATATGTTA AAAC TAGTAT ACTAGCAACT GCATGCTTGA CGTACGGATT 360
 GTCCGGGTCA TCCTTGCCGT TCCAGGTATG GATCTCAAGG TCCTCCCTGT CACGCAGATC 420
 ATTGTGGCTG ACTACCCCGA CATGGGGGTC CCGGCGAAAG AACGACCGCA CCCTGCAACG 480
 TCACTTAGTG GCTGTCTGAC AGGGATTGAT TGGCAATGAT TGGCAGACTA ACCGTCCCAG 540
 AAACGAATCA TCGAATCAT TGGTGTGCAC TGTGCTAGCG CGGCGCAGCG ACGGGCGTGG 600
 CTCCTCGGCA GGGGTGGACA TGACTTTCTC TGTGTAGATT CCTTGCAAGG TATCTGCTTG 660
 ACAGATGCAG AGATGATGTT GAGTAAAATA GGGTGCAGAG ACCCTGCAGA ACCTGCATCG 720
 GCCGGCGGGG ACATGTCGTC ATGACCGACC CGGGCCGAGC GGAATTAATG AAAC TCGGAA 780
 CAGCCAGGGG AATCCGACGC TTTGAGTGGC CAATCCATGT CTATCCACAT CTGCATTGAA 840
 GTGGATGAAA GGTGGAGGAG GGCTGGTTGC CAAATTCTGT TAGGGCTTGG AATGTAAGCA 900
 ACCGGTTTCG GTCATGACAT CATCACAGCG ATGACTTCAT CACTCTCGGA CGAGCATATA 960
 TAGTTGTGCC TGTCGTCGTA TCTCAACAAA CATCAACAAC AACAACAATC ATCTACTGCA 1020
 TTTACCAAAC TCATCTCTAC CTCAATCAAC 1050

<210> 16
 <211> 1032
 <212> DNA
 <213> *Aspergillus niger*

<400> 16
 TTCTCACAAG CAAATCCGAG AAACGCAATT GACCATATGT CCTTAGTATG AGGGCTAGCA 60
 GGAATACTTG TTTGTCGATG CATACCGACT AATAAAAATG GTTCTTCTTT TTCGCGCAGG 120
 GATTTGGTTT TTCATTTTTC ATTTTTCATA TTTTATGTTA TTTTATTTTA TTTT TTTGTT 180
 TTATTTTATT TTATTTTTC TTTT TAAATA AGACTGGGTG TTTTCGTCTC CCACCCCTGC 240
 TGAGGAAC TGTTGATATAC AGAAGCCTGA GATACTTGAG ATTATGTGAC AAAGAAATTG 300
 TAAACCTGAA TGTGACTAAG GCAGAAGAGA GCAGGTGAGG CCACACAATA CTGTACACCC 360
 ACTTCCGATC GTCATCGAAA TGAGCGTTGT GGGTACTGGA ACGACAAACC ATAGATCGGA 420
 TCTGCATCCG AATAGGCTTT ACTTTGTACA TTGGTAGGAG TTTGAATGTT TGAAGATCCG 480
 GCAGATGATC ACCTGCTGCA GGAATAATT TGAAACCTGT CCGTCTGCTT TGACCCAATT 540
 GACGAGGTAC GTGATTTGAT GCTACCCAAG CAGTATGAAT GCGACGGAGG TGGCTATCCT 600
 TACCTTGTC TAATGCAACTC ATAACGGGTA GTCCCCATGG AGCCACGGCT ATTTATTTCT 660
 AAGGGCCAAC AGTGAAAGAT TGTGAGATCG TATACCTCCT GGGTATGACG CAACCAGGGT 720
 GAGCACTGGC ATTGCTCCAA TCCATATGGT GGAGGTATGT CTCCGAGACC ATGACGCATG 780
 TAGACCCTGA CCTGTCCCTG ATCCAGGGAT CTGGTTTGTG ACTTGGCGTA TTTTAATTAA 840
 CTGCGGGTCA GGGTCTTGCC GTCCAGACCC AATGGATGAA TTGCCTGCGT GGCCCTGCAG 900

TTCCAGCGAT	CGGCACGGCA	CGGCTATGAC	GTCGGTGGGT	GTTCCGTTTG	ATATTTAACC	960
ATAGAAGACA	GCCTTGTTAT	GCTGGTTCCT	GCCGCAAAAT	AAAGGTCTAT	CCTTACCACG	1020
TTAAGTGTCA	AA					1032

<210> 17

<211> 1035

<212> DNA

<213> *Aspergillus niger*

<400> 17

TGGTGCCTGT	GAACGAGGTC	ACCACTGATA	AATTGTCTGG	ATCCGTAACG	TCTGTATCAG	60
TGGATGGTAC	TGAAAAGTTC	CAAAGATTGA	GGGCAGCATA	TGTAGCACAT	GGTGGTAGCT	120
CTAAGTGGTT	CGAAACTTCG	ATGAATGGGC	GTGATATACA	GGGGGGTAGC	ACCTGTATAG	180
TTGATGTCAG	CGAAGGCACG	GGCCATGCAG	TGATAGACAA	GAAAAGAGGG	ATTGAATAAA	240
AGACTCGATT	CTGAAGGGGA	GTAATACATC	TTTCGGCTTC	TCACCACCCC	AGATGTAGGC	300
GTGTGTAAGG	TAAGCAAGCA	CAACATATGC	TCTCCGCCAC	TCAGGTTCCC	CCTTCAACTT	360
CTTCGTAGAG	AGTATACCCA	GGTCATCGAC	AGACCTCCTG	ATAGTACCAG	ACTGAATCAG	420
GCCAGGCAAG	TCCTGAGCGA	TTTCCTCCCA	CGGGGAATAG	TAAGGGTCTT	CCAATCTCCG	480
TACGGGAGGA	ACTTCCGGGA	GGAACCCGTT	CTGGAGGGAG	ACACCGTACG	CAACAAGTGA	540
AATGTCTTGT	GTATAAAGCA	TGATAATATA	TGCACTTCAG	TAGACAACCT	GATATTCAAG	600
CCTCCTTGAA	ATGATCCTCA	TGCAGGTAAG	AGTAGATCCA	AGTGTTTATA	TTGTATTCCC	660
AAGGAGGGAC	GGAAGAAAAC	CGGATCCAGA	GGGCTCGAGA	TAGGGCCACC	TGCATGAGCA	720
AAGACGGGTC	ATTTCCCGTT	TCACAACGAC	CATTCTCAGG	CAGAAAGGCA	GTGGATTAGC	780
GGAAAACCTA	GCAACCATCC	GGATTTCTGG	AGTGGATGCT	TTGTTATCTC	ACTTCCGGCA	840
ACACGGATGG	TGAGGCGGAA	GCGGAACGGT	ATCCGGATCC	GGAGGGAGGG	ACAAACCGAA	900
GTGCCTTGTG	TCACACGCCG	GTCCATATAT	AAATGTTAAT	TGGTGACGGT	CAGCGGATTG	960
TTTCTGGTGT	TCATCCATTT	TTTCTTGATA	TCCTGGACCC	AGTTATTGCA	CTTTAATTAC	1020
TATCCAAAAT	CCACA					1035

<210> 18

<211> 1071

<212> DNA

<213> *Aspergillus niger*

<400> 18

GTCCACACTC	TAATTGGGAT	GAGCATTGCC	TTTCAGGAAC	TTGACTTTTT	GTGGTTGCAT	60
CTTTTGTTCG	CTTACCCACG	CAGGCTTCGT	TTATGAGATA	GCGATTTTCAG	GCCTCAAATT	120
GCTGTCATCA	CTTCCGGCCC	CAAACCTCGT	GGCGGTGCCC	GCCCATCGCT	GGCAAGAAGC	180
ACAGACTAAC	GCCTTACTTC	GCCTAGTTGC	ACACGATGCA	TCAAGACTTC	CTCCTATTTT	240
ACAACGACCA	GCTCCTTGCC	ATTTGTTGTT	TCGAATGCCC	CCTTCCCCGT	GCCTTGCAAT	300
TCTAATTCCC	ATCTATACCC	TTTTGCGCTC	TCACGCCCCG	CTGACTATCA	ATTTTTCTGC	360
CAGCATAACC	CTGCGTGCAG	AGCGGTTGAC	AGCAGGCTTG	CTAAGAGTCT	TCGTCTAGAA	420
GAAGTGTCTG	GGTGGATTCC	TTTTTCGGATT	AACTTTTGCT	ATTTTAAAAG	CTGATTGCAC	480
TACGGAGTAC	ACCCGCAATC	TGATTCTCAC	TGATACTCCA	ATTAAAGGCC	ACTAATTGCT	540
GACATTTGTA	ACATATTTGC	GTTTGCTTTT	CGGTTGAAAT	TGGCGGCTTT	GTGCTTTTGA	600
TCGCTAGCAG	TCTCTGTGCA	TGTACAACCT	ATCGGCTATT	CAGTCCTATT	CGGAAGCGGC	660
ACTGAGAGGC	GGAGACAAGC	CACAGATACG	ACGAGTCTGA	TTTTGAGACG	GGCGCATCCA	720
ACGAAAGCCT	GCGATTGCAA	TTGAAGATCT	CAATTGGGGG	ACTGACAACC	TAATACCTAA	780
ACATATCGAG	GATTAGGCGC	CAACACGAGC	ATGGGAAGCT	TTGGTGATG	CGACGGCTGA	840
TCCATTTCTA	ACAACCGTTT	TGCGCGCGCG	TCCGAGGTAA	GGCTCCCGAT	CCCTTGCGCG	900
TTTGATGTGT	TCATCCGAT	GCATCTACTG	TTTATATGAG	TGCGGTGGGC	AAGCTGTGCC	960
CACCGACTGA	GCATTCTCAA	CCGGCCCAAT	GCTGTGGGAT	CGATCCTCCC	CTCCAAAGTT	1020
GTCGCTCGGT	TACACTTGAT	CGGAGCTTGT	TCTGTTATAC	ACCAAACCAT	C	1071

<210> 19
 <211> 1116
 <212> DNA
 <213> *Aspergillus niger*

<400> 19
 GCCTTTGGAG AATGTAATAT CCCTTGATCA CCCCAGACTC ACATCCAGAT TCTGACTGAC 60
 CTCGTGCATA GCCATATTTA ACTTGGTCCC TGATGATCAA AACAGTACAT GGAAGTGGGG 120
 CCGAACATTT CCCTAGCCAA AGGAGCATCA GGGATTCAAC GTGATCAGTA CAGCCTTGTA 180
 GCTCGCTGAA ACTGCCCTGG AGCCAAGCCA AAGCCTAGAC AGCAAGGTGC ATGGGCGAGC 240
 TTTGCGAGGA TTTAGTTGTT ATTTAGTATG CCAAGATTAA GCCACGCAGA ATGGCTTATG 300
 GCCATAGTAG CAACAACGGC CACGGAACGT TCTTCCATGT GGTATGGGGA AAGCCGCCGG 360
 ATATTTGGGG AGGGCGTGAT TGTTCGGTG GCTATCCTTC CGGTGGAGAT AACGGCGGTG 420
 GACCATCAAA ACTCCTCAGC CCGGATTGCT CGGACTCCGC ATCGTCAAGT ATCGGTTTCA 480
 TGACAGCCAT TCGACTACTT AGAAGACACA AATAGCTATC CAACTTACCC GGACCTTTGA 540
 TAATAATTTT CAGCAGATCT GTCGATCTCA GAGACAATAA TGCCACCGTT CTGTGTGGAA 600
 GGCAGTGC GC GTGGTTTAA CAATAATAAA ACTCCAAATT GCCCAAAAGC TAGAGCGAAG 660
 GAAAGTCGAG AAGCCCAGTC TAGACTGGCG GGCTGATCGT CCGGTTCACA ACCTGATTGA 720
 CAAATCATCT GTGACCCTCG ATGTTCCGCG GCGTTGAGCG GGTGAGAATG TGGATCCTAG 780
 TGGGGAAGAC CCAATTGCCG CCGGAAAAGG TGAGCTTG TG CACATGT CCGGTGGCTC 840
 AATTCAGGGC CATGCCAGGA CATTGCTCTT CAAAGGGATG CGATTGCTAT TCCCTATTTA 900
 CTTCTTTAAC CCGCTGACGG GCCTGTCTGG ACCGAGAATT CTTGGTCAAG TTTCAGTGGA 960
 GCGATCCAGG AAAGAACATC CGATATGGAC AGTTGCCCTC GATCTCAGGT CTGTAGTCGT 1020
 GTGACACAGA GGCCTCTACT CGCTTGCTT GCATGTCGGA GCTCAAAGGT GCGATGTGAA 1080
 TTAACACGAG GAGCGACTGT CTGTCGCAGA TGCAAC 1116

<210> 20
 <211> 3191
 <212> DNA
 <213> *Aspergillus niger*

<400> 20
 GGAGTTACGT CCCAGAGATG CTGCGGACAA TGCCGATGCG GCATACTTCT CATAACGCATC 60
 CGTCAAGTAA TTCACAACAG CCATGTAGAT ACTGTAGATG CCAATACCGG TAGCCCCGAT 120
 GCCGAGGGTC GGGACAATCC AAGGCACAGA AGGAAAAC TG GCCCAACCAT ACCAGAACAG 180
 CCCGCTGCG AACAGCAAGC TGCCAATGAT ACTAGTATAC AGACGGGCCT CCGGGATGGG 240
 TTTCCCGGGC TTCTCTTTGT TCGTTTTGCG CGACCGCAAA TACAGCATAT CCTGAAGAGG 300
 ATTGACGAGA GTACCGATTA GAGCACCAC GGAGATGGCC AATTGAATGC ACCCAGTTTG 360
 CAAAGTATTC ATGCCGTAGT TGGTGGAGAA GGTCTGAACG AACTGCTGA AGAAGAGATA 420
 CAGGATACCC CAGGCGAAAG AGATCCAGAG GGTGAAGAAG GCAACGACGG GCTCGGTGAG 480
 GAGCATTCTG GTGGGCCGTT CGAAGGAGAT CTGCATCAGA CGCCAGACGC TGGTGTGTC 540
 TAGCTCGGCT TCGGCGTAGA TGGGCGCTG CTTTTCTTTG CGGAGCTTCT TGGCCGACG 600
 GGTGAGGATC ACGTCCGCGC GGGTTTTGTA CAGGATGAAC CAGAAGATGG GTAGGAGCCC 660
 TGTTAGATAG ATAATTTGGA TGTAGAAGAT CCAGCGCCAC GGGGCGGTTT TGTGGATGGC 720
 TACGATGGCG GATCCGATGA ATGGGCCAG GGCAATTCCA ACCACACTGG TAAAACCGAA 780
 GAGCGACATA GGGAGACTCC TCGCTTTGTC ACCATACCAG ACATCACTGA TGCTACCGCC 840
 GACGATGTTG ATCGATACGG AGGACGCACC ACCACCGAAG AAACGCGTCA CTACGAGAGT 900
 AGCGTAGTTT TGGGCGAAGG CGGAGGGGAA TAGCGAGATG ATGAGAATAA TGTAGGCGAC 960
 AAAGTAGCCG GGCATGCGAC CGGAGGACTC GGTCAGAGGC ACGAATATGA GAGGCCAGAA 1020
 CGCCGCACCC ATGTTCCAAG ATGTTGTGGC CCAGTAAAGG TTAGGAAATG GGGTGTGTTT 1080
 AACGTGGAAT CGTTCTTCCA TCCAGTTGTT GCCGGAGCCG TAGGAACCGG CCGGGATACC 1140
 AGTGAGGATG GAACTATACC TATTAGCACT GCCATCTTCG AGAACAAGGC TAAGACATAC 1200
 ATGAAACAAA CCGTGACCGT GAGAAGCCAT TTATAAGAGG TGCTCCAATT GAACCTGAAT 1260
 ATGTTAAAAC TAGTATACTA GCAACTGCAT GCTTGACGTA CGGATTGTCC GGGTCATCCT 1320
 TGCCGTTCCA GGTATGGATC TCAAGGTCCT CCCTGTCACG CAGATCATTG TGGCTGACTA 1380
 CCCCACATG GGGGTCCCGG CGAAAGAACG ACCGCACCCT GCAACGTCAC TTAGTGCTG 1440

TCTGACAGGG	ATTGATTGGC	AATGATTGGC	AGACTAACCG	TCCCAGAAAC	GAATCATCGA	1500
ACTCATTGGT	GTGCACTGTG	CTAGCGCGGC	GCAGCGACGG	GCGTGGCTCC	TCGGCAGGGG	1560
TGGACATGAC	TTTCTCTGTG	TAGATTCCCT	GCAAGGTATC	TGCTTGACAG	ATGCAGAGAT	1620
GATGTTGAGT	AAAATAGGGT	GCAGAGACCC	TGCAGAACCT	GCATCGGCCG	GCGGCGACAT	1680
GTCGTCATGA	CCGACCCGGG	CCGAGCGGAA	TTAATGAAAC	TCGGAACAGC	CAGGGGAATC	1740
CGACGCTTTG	AGTGGCCAAT	CCATGTCTAT	CCACATCTGC	ATTGAAGTGG	ATGAAAGGTG	1800
GAGGAGGGCT	GGTTGCCAAA	TTCTGTTAGG	GCTTGGAATG	TAAGCAACCG	GTTTCGGTCA	1860
TGACATCATC	ACAGCGATGA	CTTCATCACT	CTCGGACGAG	CATATATAGT	TGTGCCTGTC	1920
GTCGTATCTC	AACAAACATC	AACAACAACA	ACAATCATCT	ACTGCATTTA	CCAAACTCAT	1980
CTCTACCTCA	ATCAACATGC	CTATCTCCAT	TCCCTCCGCC	TCCAGCGTCC	ACGACCTCTT	2040
CAGCCTGAAG	GGCAAGGTCG	TCGTTATCAG	CGGCGCTTCC	GGCCCTCGCG	GCATGGGCAT	2100
TGAAGCCGCG	CGTGGCTGCG	CCGAAATGGG	CGCCAACATC	GCCCTCACCT	ACTCCTCTCG	2160
TCCTCAGGGT	GGTGAGAAGA	ACGCCGAAGA	ACTTCGCAAC	ACCTACGGCG	TCAAGGCCAA	2220
GGCCTACCAG	TGCAACGTGG	GCGACTGGAA	CAGCGTCAAG	AAGCTCGTGG	ACGACGTGCT	2280
GGCCGAGTTT	GGCCAGATTG	ACGCCTTCAT	TGCCAACGCT	GGCAAGACAG	CCAGCAGTGG	2340
TATCCTGGAT	GGTTCCGTTG	AGGACTGGGA	AGAGGTCATC	CAGACAGACC	TGACGGGTAC	2400
CTTCCACTGC	GCCAAGGCGG	TGGGCCCGCA	CTTCAAGCAG	CGCGGAACGG	GCAGCTTCAT	2460
CATCACCTCC	AGCATGTCGG	GTCACATTGC	CAACTTCCCG	CAGGAGCAGA	CGTCGTACAA	2520
CGTTGCCAAG	GCAGGCTGCA	TCCATATGGC	CCGGTCATTG	GCCAACGAGT	GGAGGGACTT	2580
TGCCCCGCGT	AACAGCATCT	CCCCGGGTTA	CATCGACACG	GGGCTGTCGG	ACTTTGTGGA	2640
TAAGAAGACT	CAGGATCTGT	GGATGTCGAT	GATCCCCATG	GGACGCAACG	GTGATGCCAA	2700
GGAGCTGAAG	GGTGACATAT	TCTATCTCGC	CAGTGATGCC	AGCACATACA	CGACGGGTGC	2760
CGATTTGGTC	ATTGACGGAG	GATACACCGT	GCGGTAAATT	ATCATCCGCG	GTCGGATAGA	2820
AAATAATGTT	TATGACATGT	ATGTAATGAA	TATGAATGTT	CAACAATAGT	CTTCAATTCT	2880
TCACTTGCTA	GTATCCTGAG	TATCTTGTCG	TATAGCTCAA	CAACCTGGCA	TGGCCGGGCG	2940
CGCAGTGTAG	CTCTCCCAAG	TCTATGGATA	CAACTTGCGT	CAGTGTGGAT	CTTTGGCTCA	3000
GATTTCGGAG	GGCAGTGGGG	TAAACAAGAA	CACGCAGCCA	CTCTCGGCTT	TGCGGATCCC	3060
CACGGAACGC	CAGTCAACCC	TGAAGATTCA	GTCCGGGCCCT	GGGCTCGGAC	TGATTTGGTG	3120
AGGGGATTCT	GCATCCGTCC	AAACCTTATC	ATCGTTGGTG	AGACGGACTC	CGGACAGCCA	3180
GGAGCCTTAT	T					3191

<210> 21

<211> 4031

<212> DNA

<213> *Aspergillus niger*

<400> 21

GGGGTCAAAG	GACACAATGT	CGAAATAGCG	AGATTGCTGT	GCATCATCGG	GGCTGTCTAC	60
AGTCCGCTGG	ATTTCTTTCC	CTCGGTTAAG	TATTTCTGTC	ACGCCGGACC	CGCCAGGGCC	120
ACCTGTATTG	CTAGTAAGCG	ACCGTGGTGG	GGTGACAGAA	TGGTGGCGGT	GGTGGCATTG	180
CTGCCTATAC	CTGGGTTCAC	GATTATGGAG	CCACCGTAAC	GCTCGTCGGC	AGGATCTACC	240
TGTGCAGGCA	GTTTTATGAC	GGCCAGGGCA	GCCGTTTTAT	TTTCAATGGC	ACTAGCTGAC	300
CAGTCTAGGG	GAACCTCGAG	ACGTGCACAT	AGATAGTCCT	GGTAGCAGGG	GATGAAATTT	360
AGTGCTGAGC	CTGGCGCAGT	CTGGTAGGAG	CGGGTTAGAC	TGTGAGAAAT	TATAATAGGG	420
CGGATGAGAT	CAGTACGCTA	CTCCAGGAGA	AGTCCCCGGC	TTGCACCGCA	GACGGCATCA	480
TGGGTTGGTG	TCGCAGGGAG	GAGGGTGAGT	TGAGTAAGAG	AGAATTAATG	AAAAGGATGG	540
AGTAGCCGAC	CATATAAGTA	CTGGATGAAG	TGAGCGAGTT	AGCGTCAGCC	ACGACGGGCT	600
GACATGCATG	ATCGGGGCAC	CGCCGGCAAC	CGATACAGCC	TCAGCTCGAC	GGGAAAACGT	660
GGTTACCTGG	ACTTAGAAGA	TGATAATACA	ATGAAAAGAT	AGCGTAGATT	GCCATAAATA	720
CGCCACAATC	TATTTATTAG	GCAACAATCA	AGTAAGAGCA	GATTTGGCGA	CTCATAGTAG	780
TCCCCATGGA	GGGGCGCGAT	GGTTTTCTTC	GCCGAAATAA	TCGCTTAATG	CCTCTAACGA	840
AGATCTCAGA	ACCTCCAAAC	TGTAGTATTC	TTCTCACAAAG	CAAATCCGAG	AAACGCAATT	900
GACCATATGT	CCTTAGTATG	AGGGCTAGCA	GGAATACTTG	TTTGTGATG	CATACCGACT	960
AATAAAAATG	GTTCTTCTTT	TTCGCGCAGG	GATTTGGTTT	TTCATTTTTT	ATTTTTTATA	1020
TTTTATGTTA	TTTTATTTTA	TTTTTTTGTT	TTATTTTATT	TTATTTTTTT	TTTTTAAATA	1080
AGACTGGGTG	TTTTCGTCTC	CCACCCCTGC	TGAGGAACTG	TTTGATATAC	AGAAGCCTGA	1140

GATACTTGAG	ATTATGTGAC	AAAGAAATTG	TAAACCTGAA	TGTGACTAAG	GCAGAAGAGA	1200
GCAGGTGAGG	CCACACAATA	CTGTACACCC	ACTTCCGATC	GTCATCGAAA	TGAGCGTTGT	1260
GGGTACTGGA	ACGACAAACC	ATAGATCGGA	TCTGCATCCG	AATAGGCTTT	ACTTTGTACA	1320
TTGGTAGGAG	TTTGAATGTT	TGAAGATCCG	GCAGATGATC	ACCTGCTGCA	GGAATAAATT	1380
TGAAACCTGT	CCGTCTGCTT	TGACCCAATT	GACGAGGTAC	GTGATTTGAT	GCTACCCAAG	1440
CAGTATGAAT	GCGACGGAGG	TGGCTATCCT	TACCTTGTC	TATGCAACTC	ATAACGGGTA	1500
GTCCCCATGG	AGCCACGGCT	ATTTATTTCT	AAGGGCCAAC	AGTGAAAGAT	TGTGAGATCG	1560
TATACCTCCT	GGGTATGACG	CAACCAGGGT	GAGCACTGGC	ATTGCTCCAA	TCCATATGGT	1620
GGAGGTATGT	CTCCGAGACC	ATGACGCATG	TAGACCCTGA	CCTGTCCCTG	ATCCAGGGAT	1680
CTGGTTTGTC	ACTTGGCGTA	TTTTAATTAA	CTGCGGGTCA	GGGTCTTGCC	GTCCAGACCC	1740
AATGGATGAA	TTGCCTGCGT	GGCCCTGCAG	TTCCAGCGAT	CGGCACGGCA	CGGCTATGAC	1800
GTCGGTGGGT	GTTCCGTTTG	ATATTTAACC	ATAGAAGACA	GCCTTGTTAT	GCTGGTTCTT	1860
GCCGCAAAAT	AAAGGTCTAT	CCTTACCACG	TTAAGTGTCA	AAATGAGTGG	GCTAAGCTTT	1920
GCCCCGTCAGG	CCTTACTCCA	GGCTGCTCTT	CTTTCGATTG	TCTCCGCAAT	CCCGGCGCCC	1980
ACGGCCTTCG	TCAACAATGT	CGCTCCGCCC	GAGCCCATCA	TCACCCCTTC	GCCGGTTCAA	2040
CATCGACCTT	CTCGAGTCGC	TGGTCGAAAC	ATTCTGAGTG	ACGTCGATTC	TGATATTAAC	2100
AGCATCCTTT	CTGGCCTTGG	ATCCGACCTA	CCCTCCTGGG	TCGCATCGGG	TGTGCCCAAC	2160
TACTTCCAGG	GTTTCCCTAC	TGGGGATGCA	GTGGTGAGCT	CCTTGGGCTT	GAACAGCGCT	2220
GAGTTGGCAG	CCTTGCCAC	CAATGTCTTG	AATATCGACC	CATATGCCAA	CTGGACTAGC	2280
TCTGGCTGGA	ACGTTGCTT	CCACGGAAAC	GTCTACAAGC	AGCCCAACAC	CTCCATCTCC	2340
GACCTCAATG	ATTTGGCTGA	TGTTTTCTC	GGCAATACAA	GCATCTCCGA	CCTTTCCGAA	2400
TCCGAGCAGA	AACAGGCTCG	GAATTTAACA	GCGGAAATAT	TAGTGGTCCA	GCAAGCTCAC	2460
GTCGCCGTAA	ACACAATCCA	CCTGGAGCCT	GCACCGAGTC	AGGGATCCAG	CGGACAGTCA	2520
GGCGGCGGTG	GATCCTCTAA	TACCACTGGC	GGCACCCAAG	ACCTCACTCT	GCCCTACAAC	2580
ACTACAGTTG	AAGGCGATTT	TGACACTTTC	GTACCGATTA	GCAGCAATGG	CTTGACGGCA	2640
GGTAATGAAA	CTTCGGCGAT	ACAACGGCTT	AACGTCCATG	TGGAGGGCGC	GACAATAGGA	2700
AACAGTACTG	CCTATCTCGT	ACCTCCGACC	GGGCTCACTG	TCGTTTCGGA	CATCGATGAT	2760
ATCCTACGCG	TCACCAAAAT	CTATGAACCA	GAGCAAGGTT	TACTCAACTC	ATTTGCGCGC	2820
CCCTTCAGGC	CGTGGGAGAA	TATGCCGGAC	ATCTATCGCA	ACTGGTCGAT	CAGTCTCCCC	2880
AACCTGCAC	TTCACTACCT	AGTAAGTCTG	CTGATTGCC	CTGATTGCAT	TGTTGTGCA	2940
TACTCACCAC	TACACTACAG	ACGACGACCC	CCGAACAAGT	CACCAGGAAC	TACATGCAAT	3000
TCATCTACGA	CAACTACCCC	GGTGGATCTT	TCGACACCCG	TCCCCTCAAC	TTTAGCGACG	3060
TTTCCGCCAC	CTTGTCGATC	CGGAAATTCC	TCCTGCAAAA	GGTCTTCGAG	ACCTTTCCGC	3120
AGCGCAAGTT	CATCCTCATC	GCCGACACCA	GCAACAGCGA	CGTCATGCGT	GACTATCCTG	3180
AAATGGCGAC	CGACTTCCCT	GGTCAAGTGC	AATGTATTTT	TCTCCGGAAC	ACCAGCGCCA	3240
CCGACTCAGG	GGACAAATTC	CCATACGATA	CCTCTGGGTT	TAAGAAGCTC	AATCAGTCAA	3300
ACTACATGTT	CTTCTTGCA	CCCGATGACT	TGACGAATCT	AGACATCGCG	AATGGCCAGT	3360
GTTACAACAC	ATCCATTCCG	CAGAATTTAA	CGTTTAGTTA	TCAGGGTTTG	CCATTGGGGC	3420
TGGGTGATGA	ACCCACAGCT	GTCAATGGGT	CTGCCAATCA	TACTGCTGAG	TCGGCGGCCG	3480
GTTTGTCGGT	AAAGGGTGGG	ACTGATATGC	AGGGTTTGAT	CTTTCTGTTC	ACCCTGGTTA	3540
CGGCCTATTC	CATTTTCTTG	TAAATATCAG	TTACGATGTA	TGAGAGATAA	CGACTTTGAT	3600
ATACTCACGG	CTCAATTTAC	GTTCCAAAAC	CGAATACAGC	GGATATCTAT	GCTCGATGCA	3660
AGGACTACCA	CTATGTTACA	CTACTAAGGG	TTTATCTACT	ATGGGAAATA	TTAGTGCTTG	3720
GTTTGTTACA	GCTCACTCGA	GCTTGCGTTA	CTAACCCTT	TTATCGGTCC	ATGGATCCAC	3780
CGCCCAACGC	AGTTATCCCA	GGAGGTGGCC	GCAATGGCCA	GTCAAGTCAA	TTAAATCTCC	3840
TCTGCTAGGA	ATTGCATGGG	TTTTGACAGA	CAGAACTCGG	GGAAAGTATC	GATATTCAAT	3900
GATCTGACGG	TGGCTCCGAA	ATTCGCTAGG	CTAGCCCTTC	ATCACTAGTA	ATATTAGTTG	3960
TCCAACCTTC	GCGCTTTACA	CTTTCCATAG	CAAACCCCA	ATTCCCAGCC	AAATCCTCAT	4020
ATCCTCTAGG	A					4031

<210> 22

<211> 3771

<212> DNA

<213> *Aspergillus niger*

<400> 22

GCTTCTATGG	TATCAAGGGT	CGCGTAGCCA	AAAGTTTAAAC	TGCATCGCTG	CGGTATTAGT	60
CGGTGTAGCT	GGCTGCAGCT	TTCGTTGTGT	CCCTGGTCTG	TTTAAGAAAC	GGTATCAGAT	120
CTGTCCCGCC	CGTCCCCCTG	ACACCGGCAT	TGAGTTTCTC	ATGAGAATTG	TCCACTATTG	180
TCGACGTTGT	TGTGGCTAGG	TTTGTTCCTG	AAGATACTTG	GTTTGGCGGC	GGAGTCCTTG	240
CAGCCATGAT	GATGTATCGA	GTCACCATCC	GCACATGCTT	ATCGCGAAAT	CCCCCAAGTG	300
ACATAACCGC	CATGTTATAT	GCATCCCTCA	AAGCAGATCC	GGGCTTGTGA	CTTAATACAT	360
AGGATCTCAC	ATTAGACGTG	CGGAGCATCA	TTTCCAAGAA	GCGTCGGTGA	GGGCCTGGCA	420
TGTAGTTTCG	CATGTTCTTG	AATATACAGT	CAGTATCGAC	ATTGCTTTGT	CTGGATAACT	480
CCTCGAAGAT	CAGTCCTTAC	CTGCATAAAA	CCAGTCATGA	CTTGAGGATT	CGACTTGCCG	540
TGAGGCTTGG	TCTCCCCAGT	GGCCGAATGT	TCCACACCCA	AGAAAATGTC	GAAGGCTTGG	600
ATCAGAGAAC	TTTGAGCATT	GCTACCACCA	CTGTATTGAT	GCCATTTCGCC	TCGTCTCTCC	660
CCAACATCAT	AGAATACGCC	GTTTGGCAGC	CCGGCCGATG	CCATGTTCTT	GCTGCCGGCT	720
AGGTATGGAC	GTAATTGATG	GTAAAAGACG	GGCGGCGCAC	ACTTTTCATA	CATGCGTTCC	780
AAGAGGTCAC	CGAGTTCCTT	GAGACCATAA	TTGAGCTGCT	CCAGACGAGA	AATAACACGC	840
TGGCTATCAT	CAACATTAGC	TGCATGCATA	GCATCCAGCA	TTAGCTCGAT	GAGCTGTGCT	900
CCCTTGGCCT	CGATAGCGAC	GGAGAGGACG	AAAAACCATT	CTTCGTCTTT	GGTGCCTGTG	960
AACGAGGTCA	CCACTGATAA	ATTGTCTGGA	TCCGTAACGT	CTGTATCAGT	GGATGGTACT	1020
GAAAAGTTCC	AAAGATTGAG	GGCAGCATAT	GTAGCACATG	GTGGTAGCTC	TAAGTGGTTC	1080
GAAACTTCGA	TGAATGGGCG	TGATATACAG	GGGGGTAGCA	CCTGTATAGT	TGATGTCAGC	1140
GAAGGCACGG	GCCATGCAGT	GATAGACAAG	AAAAGAGGGA	TTGAATAAAA	GACTCGATTTC	1200
TGAAGGGGAG	TAATACATCT	TTCGGCTTCT	CACCACCCCA	GATGTAGGCG	TGTGTAAGGT	1260
AAGCAAGCAC	AACATATGCT	CTCCGCCACT	CAGGTTCCCC	CTTCAACTTC	TTCGTAGAGA	1320
GTATACCCAG	GTCATCGACA	GACCTCCTGA	TAGTACCAGA	CTGAATCAGG	CCAGGCAAGT	1380
CCTGAGCGAT	TTCTCTCCAC	GGGGAATAGT	AAGGGTCTTC	CAATCTCCGT	ACGGGAGGAA	1440
CTTCCGGGAG	GAACCCGTTT	TGGAGGGAGA	CACCGTACGC	AACAAGTGAA	ATGTCTTGTG	1500
TATAAAGCAT	GATAATATAT	GCACTTCAGT	AGACAACTTG	ATATTCAAGC	CTCCTTGAAA	1560
TGATCCTCAT	GCAGGTAAGA	GTAGATCCAA	GTGTTTATAT	TGTATTCCCA	AGGAGGGACG	1620
GAAGAAAACC	GGATCCAGAG	GGCTCGAGAT	AGGGCCACCT	GCATGAGCAA	AGACGGGTCA	1680
TTTCCCCTTT	CACAACGACC	ATTCTCAGGC	AGAAAGGCAG	TGGATTAGCG	GAAAACCTAG	1740
CAACCATCCG	GATTTCTGGA	GTGGATGCTT	TGTTATCTCA	CTTCCGGCAA	CACGGATTGGT	1800
GAGGCGGAAG	CGGAACGGTA	TCCGGATCCG	GAGGGAGGGA	CAAACCGAAG	TGCCTTGTGT	1860
CACACGCCGG	TCCATATATA	AATGTTAATT	GGTGACGGTC	AGCGGATTGT	TTCTGGTGTG	1920
CATCCATTTT	TTCTTGATAT	CCTGGACCCA	GTTATTGCAC	TTTAATTACT	ATCCAAAATC	1980
CACAATGTCT	AAGTCGAATG	GCGTCAGCCT	GGCATTCCCC	GCCGAAGCAG	CGACAAAGGA	2040
ATATGCAGCA	TCTCTAGACT	CTTCTGATAG	ACTTGCTGCA	TTTCGGGAGA	AGTTCATCGT	2100
CCCATCGAAA	GCAAATATTG	CTTCCAAAAA	GCTAGCAAAG	CCTGGTGGGT	TCCATCTTGG	2160
CAACTGTAGA	CCTCGTCACT	AATGTCACAG	GCCTTTTCGCC	CGAGTCATGC	ATTTACTTCT	2220
GCGGTAACTC	ACTTGGCATT	CAACCCAAGG	CCACAGCAAA	GTAATTAGAA	GCACAGTTGG	2280
AACTTGGTTC	ATCTATTGGA	GTTGGCGGCC	ACTTCACTGA	TCTTGAGGGC	TCACCCTTGA	2340
AGCAATGGCA	GCTGCTCTCA	GAGCAAGCAG	CTGATTCAAT	GAGCAAGATT	GTGGGAGCAA	2400
AGCCAGAAGA	AGTCGCAGCG	ATGGGAACAC	TTACAACGAA	CCTTCATCTT	CTGCTGGCTA	2460
GTTTCTATAA	ACCGACTCAA	ACAAAGCACA	AGATCTTGAT	GGATTGGAAG	GCTTTCCCAA	2520
GTGATCACGT	AAGTTCGAAA	GAAGCTGCCT	TCTGCTATCA	CGGCACTGAC	GATTATATAG	2580
TACGCTATCG	AATCTCATAT	TGCCTGGCAT	GACCTGGATC	CTAAGGAGTC	TATGGTGCTC	2640
ATCGGTCCAG	ATGAAGGCGA	ATACGAGATA	TCCACCCAGA	AGATTTTCTC	TTATATCGAC	2700
AAGCATGCTG	ATGAGGCTGC	TATGATTCTT	CTTCCTGGTA	TTCAATATTA	TACTGGGCAA	2760
CTGTTTCGATA	TCCAGAAAAT	TACAAAATAT	GCCCATTTCG	GCAACATGGT	TGTTGGCTGG	2820
GATCTGGCCC	ATGCGTTCGC	TAATGTTGAG	CTAAAGCTGC	ATGATTGGAA	CGTTGATTTT	2880
GCAGCATGGT	GCACATATAA	GTACGGAAAC	GCCGGTCCTG	GAGCAATGGG	GGGGCTTTTC	2940
GTGCATGAAC	AACACGGAGA	GGTCGATTAC	AGCGCGGGAG	AAGATGCACC	CAAATTCCGC	3000
CATCGTCTGA	CCGGGTGGTA	TGGTGGCGAT	CGATCGGTAA	GATTCAAGAT	GGACAACAGT	3060
ATGTGTGCAC	TTGGCTATAG	AACAATTTTG	TGGGCTAGTA	AGCTGATACG	ATTGATAACC	3120
AGAGTTCAAA	CCTATTCTCT	GGGCAGGAGG	ATTTTCAGATA	TCAAATCCTT	CGGCCATCGA	3180
CCTTGCGTGT	CTTTGTGCCG	CTTTATCGGT	GTTTGATGAG	ACGTCAATGG	CAGATCTTCG	3240
CAGAAAATCA	TTGAAGCTAA	CAGCATACCT	TGAGTTCCTT	TTGCTCAGGG	ATTATGAAGA	3300
AGAATCTAGG	CCATTTCAGCA	TTATCACGCC	CAAGGACCCC	GAGGCGAGAG	GCGCACAAC	3360
GAGTCTATTG	CTCAAGCCTG	GTCTATTACA	GAATGTTGCA	CAGAAGTTGC	AAGAAGCAGG	3420

AATAGTCTGC	GATAAACGGG	AGCCGGGTGT	TGTGCGTGTC	GCTCCGGTTC	CTTTGTACAA	3480
CAGCTTCAGC	GAGGTTTGGG	CATTTGTGAA	GATCTTCAAG	GATGCACTGC	AGCAGTGATG	3540
GGTCTCACTA	GATCATTTGG	TCAGGACTGA	GACCAGATTT	GTCAAGTTTCG	CTCCGATGAA	3600
TGATTTTTTA	GATAGATACC	AAATTATTTG	ATTCTTAGAT	TTGTTGAAGC	TCTCACTGCG	3660
ACATGGTCCA	CAGCCACTGT	GCAAAAGCAA	ATAACTGAAG	TAGGAAGTCT	ATGATAGACC	3720
CGTACTGTTG	AGGTTATAGA	TTTCCAAAGC	CGTTTCCTTG	ATCACCACCT	C	3771

<210> 23

<211> 3562

<212> DNA

<213> *Aspergillus niger*

<400> 23

CCTGTGCTTG	TGGTATGACA	GAGCCACTGA	TATCATCTCC	TCTCATCTCC	AGTTGTAATC	60
GGAATTGATC	GAGGGGGCCA	ATCTTCCACA	CAGCCCCATG	GAAGACGCAC	AGATGCTTTC	120
GATGCCACGC	TGGCAGACGG	CCGCCAACTG	TCATCGTCCA	GTGCTGCCGC	CATTCCCATG	180
CAAAAGCCAA	CGCTAAGAAG	CGCGCGGGCG	CCGAGTTTCC	ACGCTTGCAG	CTTAGTTCAG	240
CGACCCAGCT	CAGTGGCGGC	TCGCTCATGC	ATAGAAAGAT	GGGTTTGGAC	CCGCGCTGGG	300
CCTTTCATCT	GCAACACTGC	TCGCAGACCA	TGTGTGACTT	TTGGCTGCAA	AAAAAAGAG	360
GCTGGCGACA	GCTGATCCAC	TGTGCTATTG	CTATCTCATC	CTCCCCGGCT	GGATCGAAAG	420
CTGCACGGCA	GGGGAAATCA	ACAGTTGGGC	TCTGACACGT	ATCGAGCTAA	CTGATGCGCT	480
GTCCTTCCGC	CTAATGCCTA	CTAAGTCTTA	GGGACGGGAC	CATCATCCAA	CATGCCACCT	540
TGTCCTTCGT	AATTCCTTTT	TGATCTCGGT	GATGGTGGCC	ACCCAGGAAA	TGACGTTCCCT	600
CACCTTGGCT	GCAATTGGTG	CCCGATGTCT	CATCTCCGCT	CAGAACGGCG	GCGAGGGCAG	660
AATAAAAATT	GACGGGCACA	GCAGGAACTA	ATCAAACATA	CCGGCGGTCT	TCTAAATGAT	720
CGTATCTCCC	GGAATACTCC	CTCTTTCTAC	AGAAGCGGGG	ATCCATTTCG	CCTATATGTA	780
TAGTGGAGCT	TTGCCGCACT	GATTTCGCACG	GCATTATACC	TCCGCACTGC	CGTGCCTTCA	840
TCCACTTGAT	TGTGGCCAGG	GAAATATCAA	CGGGGCAGAT	AACTTGCCCT	TGCATCATTA	900
TTGCTCCGCA	CATTCAATTG	TTATCGCTCT	GATCCCACAT	CCTCACGCAG	CTGCGACCGG	960
ATTGTCCACA	CTCTAATTGG	GATGAGCACT	GCCTTTCAGG	AACTTGACTT	TTTGTGGTTG	1020
CATCTTTTGT	TGCCTTACCC	ACGCAGGCTT	CGTTTATGAG	ATAGCGATTT	CAGGCCTCAA	1080
ATTGCTGTCA	TCACTTCCGG	CCCCAACTC	GTGGGCGGTG	CCCGCCCATC	GCTGGCAAGA	1140
AGCACAGACT	AACGCCTTAC	TTCGCCTAGT	TGCACACGAT	GCATCAAGAC	TTCTCTCTAT	1200
TTCACAACGA	CCAGCTCCTT	GCCATTTGTT	GTTTTCGAATG	CCCCCTTCCC	CGTGCCTTGC	1260
ATTTCTAATT	CCCATCTATA	CCCTTTTGCG	CTCTCACGCC	CGCCTGACTA	TCAATTTTTC	1320
TGCCAGCATA	ACCCTGCGTG	CAGAGCGGTT	GACAGCAGGC	TTGCTAAGAG	TCTTCGTCTA	1380
GAAGAACTGT	CTGGGTGGAT	TCCTTTTCGG	ATTAACCTTT	GCTATTTTAA	AAGCTGATTG	1440
CACTACGGAG	TACACCCGCA	ATCTGATTCT	CACTGATACT	CCAATTAAAG	GCCACTAATT	1500
GCTGACATTT	GTAACATATT	TGCGTTTGCT	TTCCGGTTGA	AATTGGCGGC	TTTGTGCTTT	1560
CGATCGCTAG	CAGTCTCTGT	GCATGTACAA	CTGATCGGCT	ATTCAGTCCT	ATTCGGAAGC	1620
GGCACTGAGA	GGCGGAGACA	AGCCACAGAT	ACGACGAGTC	TGATTTTGAG	ACGGGCGCAT	1680
CCAACGAAAG	CCTGCGATTG	CAATTGAAGA	TCTCAATTGG	GGGACTGACA	ACCTAATACC	1740
TAAACATATC	GAGGATTAGG	CGCCAACACG	AGCATGGGAA	GCTTTGGTGC	ATGCGACGGC	1800
TGATCCATTT	CTAACAACCG	TTTTGCGCGC	GCGTCCGAGG	TAAGGCTCCC	GATCCCTTGG	1860
CGCTTTGCAT	GTTTCACTCC	GATGCATCTA	CTGTTTATAT	GAGTGCGGTG	GGCAAGCTGT	1920
GCCACCCGAC	TGAGCATTCT	CAACCGGCCC	AATGCTGTGG	GATCGATCCT	CCCCTCCAAA	1980
GTTGTCCGTC	GGTTACACTT	GATCGGAGCT	TGTTCTGTTA	TACACCAAAC	CATCATGGAT	2040
GTGACTAATT	TGGCTGTCGA	TTGTCCCGTT	ATAGAGACCG	GTTCCCTTGC	TTTGTGGATG	2100
GCTTCCTCCG	AAGTGGACCA	GAAGGCCTTG	GGCAACTTTG	CGGCCATAAC	GGGCCTCGAG	2160
AACCCGTTCC	TGTCCGCGAT	GCTGTACAAG	ATTCTGGGGT	TGTCTGTTAT	GTTATCGAAG	2220
AAACTCCTCC	GTGCACGACG	ATTGCGGCGC	CTAGACCCCA	CCCGAGAGAC	TAAATCGCTT	2280
CATCTTTATT	ACCATATCAT	CTGGCTTTTCG	CGTGAGGGTC	TTCTGATATT	AGAGGAGTTC	2340
GTCCTGCCAC	TGGTTGAGGG	ATTTATGGAA	CTCAAAATTT	TAGCTTACAA	GCTGCGCGCC	2400
TCCTTTTACC	ATATATTCGT	CCTATTTTTCAG	AACCAGCCTG	CTGTCCACTC	GCCGGGTATC	2460
GTTAGTCTGC	CGAGTAGTAC	ACCTTTTATCC	AACGGCGTGA	CAGAAGCGGA	GTCACCATCC	2520
AAGGATTCCA	ATTTGAGATT	TTCATTTTCAG	CTGAAGCCCCG	ACACGATAAC	GGTTTCCGGG	2580

AAGCCCAGTT	CCGCCTCGGA	TAGTGCGCCA	CGAGGCAAGG	TTACGCAGGC	ACCCCCGGGC	2640
TTTGCGCCGG	TTCAACCGCC	CAAGTCGAAC	TCGGCGTTTC	TCCTCCCTGC	TCTCGATTAT	2700
ACCCCTACAG	CTACCGCATG	CTTCAATCAT	GCCGCCCTCT	TGGCAGATCA	ATTCCCTCCG	2760
GGGTCGCACC	CGCTCCGTCT	GTCCATCAAG	CTTGAATTTG	CCGCCTACCT	CTACGATTGC	2820
CTACATGACG	CTCATGCCTG	TCGACGGTTG	GCCAAGCAAG	CCATTGCCGA	TGTGTACAAT	2880
GCGCAAGAGG	GCATGGACGA	CGAGAGCTTC	GAGGATGCAG	CTGAGATTGT	GGGTATTTTG	2940
GGCAAGATGG	TGAAGCGGGG	AAGGAACACA	AGCAGTGCTG	GAAATAGCAC	CACGGCTGTG	3000
AAAACGCCCC	AGGAGGAGCG	AAGTGAGGGC	AGTCGAACAC	CATCCTCTCA	GACAACGTTG	3060
AAAAGGCCCG	CGCGAGTGCC	GAAAAGCACC	TCCTCACC GG	CGCGCGCGAC	CAAGCCGACG	3120
ACTAGCGAGG	GGATGTCCCC	CGCTGTGCCC	GACCCACCA	TGATGAATCC	CATATGATGC	3180
CTTTCCCTAC	GACCTGTATA	TGCTTCTTCT	TGGAACATATG	GAGTCGCGCA	TGAGCGCTTG	3240
GCCTGTGACT	TCTACTTGAC	TTGCACCTGT	CTGTATCATC	TCGGTAACGG	AGTTGCGGCG	3300
ATAGACTGGA	ATTGGGCATC	AGCTTGGCGG	GAGCTGTGGA	ATGAGCGATT	AAGCGCGACA	3360
AAAACCTTGAG	GTGGAAGGGA	CAGACCAGCC	AATTTGTGAT	GGTGACTGTA	GATAGATAGA	3420
TGTGTGCAAT	ACAAAGTATG	TTGTTGATGT	CATGCACCCC	ACTTGTGTAT	GCTTGTCTGG	3480
ACAATGGACC	TACACTTAGT	CGCAACAAC	GTATGTCTCA	AGCCAATGGC	CATGGGGCAG	3540
GATGGATCGA	TGGATCTCTT	GT				3562

<210> 24

<211> 4361

<212> DNA

<213> *Aspergillus niger*

<400> 24

CGACGGACTA	ACGTGCGCGG	CCCCATTTCA	CGTCTTGAAG	ACAGATGTGT	ATGCGATTGC	60
TCGCGAAATT	CTCATCCGAT	TCCGTGCGGT	TGACCTGACA	TCTGGAATAG	ACGAACAGGA	120
CTTGGAAGCT	ATCATTTCGT	TTACCGAAGA	GTCCAACGTA	CTTCTCACTG	ACTGGTGGA	180
GTCCGTCCCG	AAACTGTATG	ACTTCGATTA	CTGGAAGCA	GACGGTCGAT	GGGAAATGAT	240
TGAAGGCCAG	GTGCAAAGGC	TTCCCTCCGA	AGCAAGGCCA	CAGGTGGAGA	CAATATATCT	300
ACAGGCTGCT	TTTCTCCAGT	TGACGTACGA	CGGCATGGTC	ATCCAAGCAA	ATAGACCACT	360
GTTGGAGCGA	AGAATCAACA	GGCTAACATC	TTGCGGGGCG	ATCGTGATG	CGATGCATAG	420
GGCGCTTGAC	CGAGCCACTG	CTGCTGCTCT	CCGCATATCC	CGTGTTCCAG	TTCACAAGCT	480
GAAAAATCAT	TTTGCTATAG	CAGTCGTATC	GATGCAAGAG	TTCACTGCTG	GTGTGATTCT	540
CTGTATACCT	CCCACGGTGA	AGCCTTTAAC	GGCCACCGCT	CATGAGGCGA	AGGACGGGGT	600
GGTAAGAATA	ATACGCGCCA	GTCGAAGTCT	CAAGAGCCAC	GATCGAATAG	CAAGGCCAAC	660
TGAACAACCT	CTGACGGAAC	TACTCAAAGT	TACCACCCAA	CGCGAGATAA	CGCACGCCTT	720
GGCCGATAAT	CTCGACACCA	ATTTCCCTC	CGATACGCGG	CTAACAGGCT	CCGTTGTTGC	780
CGGCGCCTCA	AGTGACTCCT	TGGGTGCAGG	ACGCTCCAGG	CCAGTTCCCG	GATCCGCAGC	840
GCAGGCAGTT	GATGATCCTA	TCCAAGGAGC	CTGGGACTCA	GAATTCAATG	CCCTCCCTGC	900
CGTAACGGGA	CCTTATGATT	TCCTCCCTGC	TCAAGCCTTT	CAACAGCTGG	ACGATACATT	960
TGGTGCCTTT	GGAGAATGTA	ATATCCCTTG	ATCACCCAG	ACTCACATCC	AGATTCTGAC	1020
TGACCTCGTG	CATAGCCATA	TTTAACTTGG	TCCCTGATGA	TCAAAACAGT	ACATGGAACT	1080
GGGGCCGAAC	ATTTCCCTAG	CCAAAGGAGC	ATCAGGGATT	CAACGTGATC	AGTACAGCCT	1140
TGTAGCTCGC	TGAAACTGCC	CTGGAGCCAA	GCCAAAGCCT	AGACAGCAAG	GTGCATGGGC	1200
GAGCTTTGCG	AGGATTTAGT	TGTTATTTAG	TATGCCAAGA	TTAAGCCACG	CAGAATGGCT	1260
TATGGCCATA	GATGCAACAA	CGGCCACGGA	ACGTTCTTCC	ATGTGGTATG	GGGAAAGCCG	1320
CCGGATATTT	GGGGAGGGCG	TGATTGTTCC	GGTGGCTATC	TTTCCGGTGG	AGATAACGGC	1380
GGTGGACCAT	CAAAACTCCT	CAGCCCGGAT	TGCTCGGACT	CCGCATCGTC	AAGTATCGGT	1440
TTCATGACAG	CCATTCGACT	ACTTAGAAGA	CACAAATAGC	TATCCAACCT	ACCCGGACCT	1500
TTGATAATAA	TTTTACGACG	ATCTGTGAT	CTCAGAGACA	ATAATGCCAC	CGTTCTGTGT	1560
GGAAGGCACT	GCGCCGTGGT	TTAACAATAA	TAAACTCCA	AATTGCCCAA	AAGCTAGAGC	1620
GAAGGAAAGT	CGAGAAGCCC	AGTCTAGACT	GGCGGGCTGA	TCGTCCGGTT	CACAACCTGA	1680
TTGACAAATC	ATCTGTGACC	CTCGATGTTT	CGCGGCGTTG	AGCGGGTGAG	AATGTGGATC	1740
CTAGTGGGGA	AGACCCAATT	GCCGCCGGAA	AAGGTGAGCT	TGTGCTGCAC	ATGTCCGGTG	1800
GCTCAATTCA	GGGCCATGCC	AGGACATTGC	TCTTCAAAGG	GATGCGATTG	CTATTCCCTA	1860
TTTACTTCTT	TAACCCGCTG	ACGGGCCTGT	CTGGACCGAG	AATTCTTGGT	CAAGTTTCAG	1920

TGGAGCGATC	CAGGAAAGAA	CATCCGATAT	GGACAGTTGC	CCTCGATCTC	AGGTCTGTAG	1980
TCGTGTGACA	CAGAGGCCTC	TACTCGCTTG	CCTTGCAATG	CGGAGCTCAA	AGGTGCGATG	2040
TGAATTAACA	CGAGGAGCGA	CTGTCTGTCT	CAGATGCAAC	ATGAGAAATT	TGGAGTGTGT	2100
CTACACAAAA	TCACGAAGGT	ATCCCTCGTC	TATCTTTCCG	TGTGTATTAT	TAGCCGAGAG	2160
TCACGCAATG	CTCTTTGTAT	ACAATGAAAT	AATGTTGGAG	CTAGAGAATG	ATTTCTGATA	2220
GTTATATGTT	ACAGGACACT	CAGGAGGTCG	TCAACTAGTC	CGAACACGGC	GTCCTCTCCT	2280
GACACTAATG	CCTCATCTTC	GCCATGCTCG	GA CTCTAGTG	GGACCCATGC	TGGCATTTC	2340
ATTCCCCACT	CAGATAGTCC	CGCCAGTGTG	GACGAATGTG	GGAACAACGG	CCTTTTCGAC	2400
GGGGGTAGGT	AATATGTCAT	GCTTACCTGC	GACACTCACT	CACCATTCTT	CTAGAATCTC	2460
TAAC TCAATT	ATTGGTCTCT	GGGAGCAAAG	CCATATCAGG	CCAACCACAT	ATCGCCATTA	2520
AGCCACATGT	TCTGTCTACC	TATCCCCTTG	ACAACCTCCT	CTCAGACATT	GAGGGACTCC	2580
TCGAGCAGCT	AGTCTCTCAT	TTGGCTCAAA	GCAAGGAAGA	TCTAACAAAA	CCGGACCGGA	2640
TGTTTTCGAAC	TCGCCTTCCA	GCGGTCTCAG	GTGAGGACAA	TGAGTATCTG	CTAAGGAAAG	2700
GAGCTCTGCT	GCTCTTGCCA	CCAGCCTTGC	GCAACGAACT	ATTGGCGGCA	TATGTCAACT	2760
ACGTCCATCC	GCTATTACCG	ATTCTTGATC	TTGGAGACTT	TCTCTCACAA	ATCATTGTTG	2820
GAGATGATGC	TCAATTCAAG	TCACCACTAT	TATACCAAGC	TGTCATGTTT	GCGGGCAGTA	2880
TTTTTGCTCA	GCGAGACGGC	ACAGACGAAC	CTGAACAGCT	AACGAGGGCT	TTATTTGAAC	2940
GTACAAAGGT	TAGTAATGGT	TACAATTAGA	CATATCTTAG	ATATACTGAG	CGATTTCCAG	3000
GTTCTATACG	AGTTTGAATC	CGAATCATGT	GCCTACACTC	AAATTCAAGC	TCTCTTATTG	3060
ATGACTCTCT	GGCATGGAGA	TATGTCACGC	CACAAAGGCC	CGTCTTACTG	GCTTGACGTC	3120
GCCTTCTCCA	CTGCGGAAAG	AATTGGCCTA	CTCCATGATG	AAGACTGGCC	ATGCGATAGC	3180
CGTTCCTTCC	GATCCGCTAT	GTGGTGTGTC	CTTTACGTAC	GTGATCGAAC	AATCTCGCTT	3240
GGGTCTCGTC	GGCCTCCACG	GATGCCAGTC	AACAAATATC	CGGACTCAAT	TCTCCATCTT	3300
ATGAGCGACG	CACCTCGGGA	ATATGACGAA	ATAATTCTTG	AGAGTCTAGG	TAGCGCTTCT	3360
CTGATGCTTA	CAAGCACTTA	CCAAGAACAA	TCGACCATGC	TCTTCAAGGA	ACTTGTCAAA	3420
CTCTCCTACT	GCGTCGGGGC	AATCTTAGAT	TATCTATACG	AAGGGGCCTG	GGTCCGAATA	3480
CCAACACGTC	GCAGCAGCTT	TTATTCACTC	GCCCCTAAAT	GCAGCATCTC	GCCATCCATC	3540
AGACCAAGCT	GTGAGAAACT	ACTTTACTCT	TGGATTCAAT	TCCTCCCTCT	CAACGCAGTC	3600
TACCATCCGC	CACAGTTGCC	CGCCGATGCG	GAAGGGGAGT	CAGCCGAAAC	AGTCTTCCTT	3660
GTTCCACAAAG	CCTTCTTTTA	CTTGCTGTAC	CTAGCCTCCA	TGGCAGTCAT	ATACCGCACA	3720
GGCACTCATA	GCCAACAAAC	CAGCACGGCA	ACTCCGAACA	TGGAAGGTCT	ACGGGCATCG	3780
ACGTCACAAA	TAAAAAAGGT	CCTTGCGGAG	CTGCAGGGCA	TGGGACTGCT	TCAGTTCCTT	3840
CCAGGGGCAT	CAGTCACTAT	CCTGATGTTT	GCTGTGGAGG	CTAGTCTGTT	GGATCTCCAG	3900
GGCTCGGATA	CCATGGTCCG	AAGGCAGGCT	ATGTCAAATC	TATATGCCTG	TGAGGAGGCT	3960
GCGTTACACT	TGATGCAGGC	TTATCCTACG	GCTGAGATGG	CGGTATTGAA	GA CTAGGACA	4020
GCCCGTGCAA	ACCTTCTTGG	GTTGGTTGAG	ACATGAAGAC	CAATCATGGA	AGCTCTGCTT	4080
CCATCAAAGC	ATTGTGTCTT	TGCTTAATGA	ATGAATGGAT	TATGAGCGTT	TCGCAACCTT	4140
CTTTAATGCA	GCTACATAAT	AGAGCATAAA	CAGAGGCCCG	GAAATATATA	ATTTGGCTAA	4200
GAACCCGGTT	GCAAGCAGGC	AAGTCACGAG	GTGCAGAACA	TACCATCAAG	TTTATAGATT	4260
ATTGTAGGGG	ATATCCTCAA	CAATATCTTA	CTTCCACCAT	TACGTACTGA	AGATCGAGGT	4320
TATCGGGGAC	CATGCTGTCT	TGTCAATAGG	TACGACTCGT	T		4361

<210> 25

<211> 801

<212> DNA

<213> *Aspergillus niger*

<220>

<221> CDS

<222> (1)..(801)

<400> 5

atg cct atc tcc att ccc tcc gcc tcc agc gtc cac gac ctc ttc agc	48
Met Pro Ile Ser Ile Pro Ser Ala Ser Ser Val His Asp Leu Phe Ser	
1 5 10 15	
ctg aag ggc aag gtc gtc gtt atc acc ggc gct tcc ggc cct cgc ggc	96

```

Leu Lys Gly Lys Val Val Val Ile Thr Gly Ala Ser Gly Pro Arg Gly
      20      25      30
atg ggc att gaa gcc gcg cgt ggc tgc gcc gaa atg ggc gcc aac atc      144
Met Gly Ile Glu Ala Ala Arg Gly Cys Ala Glu Met Gly Ala Asn Ile
      35      40      45
gcc ctc acc tac tcc tct cgt cct cag ggt ggt gag aag aac gcc gaa      192
Ala Leu Thr Tyr Ser Ser Arg Pro Gln Gly Gly Glu Lys Asn Ala Glu
      50      55      60
gaa ctt cgc aac acc tac ggc gtc aag gcc aag gcc tac cag tgc aac      240
Glu Leu Arg Asn Thr Tyr Gly Val Lys Ala Lys Ala Tyr Gln Cys Asn
      65      70      75      80
gtg ggc gac tgg aac agc gtc aag aag ctc gtg gac gac gtg ctg gcc      288
Val Gly Asp Trp Asn Ser Val Lys Lys Leu Val Asp Asp Val Leu Ala
      85      90      95
gag ttt ggc cag att gac gcc ttc att gcc aac gct ggc aag aca gcc      336
Glu Phe Gly Gln Ile Asp Ala Phe Ile Ala Asn Ala Gly Lys Thr Ala
      100      105      110
agc agt ggt atc ctg gat ggt tcc gtt gag gac tgg gaa gag gtc atc      384
Ser Ser Gly Ile Leu Asp Gly Ser Val Glu Asp Trp Glu Glu Val Ile
      115      120      125
cag aca gac ctg acg ggt acc ttc cac tgc gcc aag gcg gtg ggc ccg      432
Gln Thr Asp Leu Thr Gly Thr Phe His Cys Ala Lys Ala Val Gly Pro
      130      135      140
cac ttc aag cag cgc gga acg ggc agc ttc atc atc acc tcc agc atg      480
His Phe Lys Gln Arg Gly Thr Gly Ser Phe Ile Ile Thr Ser Ser Met
      145      150      155      160
tcg ggt cac att gcc aac ttc ccg cag gag cag acg tcg tac aac gtt      528
Ser Gly His Ile Ala Asn Phe Pro Gln Glu Gln Thr Ser Tyr Asn Val
      165      170      175
gcc aag gca ggc tgc atc cat atg gcc cgg tca ttg gcc aac gag tgg      576
Ala Lys Ala Gly Cys Ile His Met Ala Arg Ser Leu Ala Asn Glu Trp
      180      185      190
agg gac ttt gcc cgc gtg aac agc atc tcc ccg ggt tac atc gac acg      624
Arg Asp Phe Ala Arg Val Asn Ser Ile Ser Pro Gly Tyr Ile Asp Thr
      195      200      205
ggg ctg tcg gac ttt gtg gat aag aag act cag gat ctg tgg atg tcg      672
Gly Leu Ser Asp Phe Val Asp Lys Lys Thr Gln Asp Leu Trp Met Ser
      210      215      220
atg atc ccc atg gga cgc aac ggt gat gcc aag gag ctg aag ggt gca      720
Met Ile Pro Met Gly Arg Asn Gly Asp Ala Lys Glu Leu Lys Gly Ala
      225      230      235      240
tat gtc tat ctc gcc agt gat gcc agc aca tac acg acg ggt gcc gat      768
Tyr Val Tyr Leu Ala Ser Asp Ala Ser Thr Tyr Thr Thr Gly Ala Asp
      245      250      255
ttg gtc att gac gga gga tac acc gtg cgg taa      801
Leu Val Ile Asp Gly Gly Tyr Thr Val Arg
      260      265

```

<210> 26

<211> 1602

<212> DNA

<213> *Aspergillus niger*

<220>

<221> CDS

<222> (1) .. (1602)

```

<400> 26
atg agt ggg gta agc ttt gcc cgt cag gcc tta ctc cag gct gct ctt      48
Met Ser Gly Val Ser Phe Ala Arg Gln Ala Leu Leu Gln Ala Ala Leu
1      5      10      15
ctt tcg att gtc tcc gca atc ccg gcg ccc acg gcc ttc gtc aac aat      96
Leu Ser Ile Val Ser Ala Ile Pro Ala Pro Thr Ala Phe Val Asn Asn
20      25      30
gtc gct ccg ccc gag ccc atc atc acc cct tcg ccg gtt caa cat cga      144
Val Ala Pro Pro Glu Pro Ile Ile Thr Pro Ser Pro Val Gln His Arg
35      40      45
cct tct cga gtc gct ggt cga aac att ctg agt gac gtc gat tct gat      192
Pro Ser Arg Val Ala Gly Arg Asn Ile Leu Ser Asp Val Asp Ser Asp
50      55      60
att aac agc atc ctt tct ggc ctt gga tcc gac cta ccc tcc tgg gtc      240
Ile Asn Ser Ile Leu Ser Gly Leu Gly Ser Asp Leu Pro Ser Trp Val
65      70      75      80
gca tcg ggt gtg ccc aac tac ttc cag ggt ttc cct act ggg gat gca      288
Ala Ser Gly Val Pro Asn Tyr Phe Gln Gly Phe Pro Thr Gly Asp Ala
85      90      95
gtg gtg agc tcc ttg ggc ttg aac agc gct gag ttg gca gcc ttg ccc      336
Val Val Ser Ser Leu Gly Leu Asn Ser Ala Glu Leu Ala Ala Leu Pro
100      105      110
acc aat gtc ttg aat atc gac cca tat gcc aac tgg act agc tct ggc      384
Thr Asn Val Leu Asn Ile Asp Pro Tyr Ala Asn Trp Thr Ser Ser Gly
115      120      125
tgg aac gtt cgc ttc cac gga aac gtc tac aag cag ccc aac acc tcc      432
Trp Asn Val Arg Phe His Gly Asn Val Tyr Lys Gln Pro Asn Thr Ser
130      135      140
atc tcc gac ctc aat gat ttg gct gat gtt ttc ctc ggc aat aca agc      480
Ile Ser Asp Leu Asn Asp Leu Ala Asp Val Phe Leu Gly Asn Thr Ser
145      150      155      160
atc tcc gac ctt tcc gaa tcc gag cag aaa cag gct cgg aat tta aca      528
Ile Ser Asp Leu Ser Glu Ser Glu Gln Lys Gln Ala Arg Asn Leu Thr
165      170      175
gcg gaa ata tta gtg gtc cag caa gct cac gtc gcc gta aac aca atc      576
Ala Glu Ile Leu Val Val Gln Gln Ala His Val Ala Val Asn Thr Ile
180      185      190
cac ctg gag cct gca ccg agt cag gga tcc agc gga cag tca ggc ggc      624
His Leu Glu Pro Ala Pro Ser Gln Gly Ser Ser Gly Gln Ser Gly Gly
195      200      205
ggg gga tcc tct aat acc act ggc ggc acc caa gac ctc act ctg ccc      672
Gly Gly Ser Ser Asn Thr Thr Gly Gly Thr Gln Asp Leu Thr Leu Pro
210      215      220
tac aac act aca gtt gaa ggc gat ttt gac act ttc gta ccg att agc      720
Tyr Asn Thr Thr Val Glu Gly Asp Phe Asp Thr Phe Val Pro Ile Ser
225      230      235      240
agc aat ggc ttg acg gca ggt aat gaa act tcg gcg ata caa cgg ctt      768
Ser Asn Gly Leu Thr Ala Gly Asn Glu Thr Ser Ala Ile Gln Arg Leu
245      250      255
aac gtc cat gtg gag ggc gcg aca ata gga aac agt act gcc tat ctc      816
Asn Val His Val Glu Gly Ala Thr Ile Gly Asn Ser Thr Ala Tyr Leu
260      265      270
gta cct ccg acc ggg ctc act gtc gtt tcg gac atc gat gat atc cta      864
Val Pro Pro Thr Gly Leu Thr Val Val Ser Asp Ile Asp Asp Ile Leu
275      280      285
cgc gtc acc aaa atc tat gaa cca gag caa ggt tta ctc aac tca ttt      912

```


Arg	Val	Thr	Lys	Ile	Tyr	Glu	Pro	Glu	Gln	Gly	Leu	Leu	Asn	Ser	Phe	
290						295					300					
gcg	cg	ccc	ttc	agg	ccg	tgg	gag	aat	atg	ccg	gac	atc	tat	cg	aac	960
Ala	Arg	Pro	Phe	Arg	Pro	Trp	Glu	Asn	Met	Pro	Asp	Ile	Tyr	Arg	Asn	
305					310					315					320	
tgg	tcg	atc	agt	ctc	ccc	aac	ctg	cac	ttt	cac	tac	cta	acg	acg	acc	1008
Trp	Ser	Ile	Ser	Leu	Pro	Asn	Leu	His	Phe	His	Tyr	Leu	Thr	Thr	Thr	
				325					330						335	
ccc	gaa	caa	gtc	acc	agg	aac	tac	atg	caa	ttc	atc	tac	gac	aac	tac	1056
Pro	Glu	Gln	Val	Thr	Arg	Asn	Tyr	Met	Gln	Phe	Ile	Tyr	Asp	Asn	Tyr	
			340					345					350			
ccc	ggt	gga	tct	ttc	gac	acc	cgt	ccc	ctc	aac	ttt	agc	gac	ggt	tcc	1104
Pro	Gly	Gly	Ser	Phe	Asp	Thr	Arg	Pro	Leu	Asn	Phe	Ser	Asp	Val	Ser	
		355					360					365				
gcc	acc	ttg	tcg	atc	cg	aaa	ttc	ctc	ctg	caa	aag	gtc	ttc	gag	acc	1152
Ala	Thr	Leu	Ser	Ile	Arg	Lys	Phe	Leu	Leu	Gln	Lys	Val	Phe	Glu	Thr	
		370				375						380				
ttt	ccg	cag	cg	aag	ttc	atc	ctc	atc	gcc	gac	acc	agc	aac	agc	gac	1200
Phe	Pro	Gln	Arg	Lys	Phe	Ile	Leu	Ile	Ala	Asp	Thr	Ser	Asn	Ser	Asp	
385					390					395					400	
gtc	atg	cgt	gac	tat	cct	gaa	atg	gcg	acc	gac	ttc	cct	ggt	caa	gtg	1248
Val	Met	Arg	Asp	Tyr	Pro	Glu	Met	Ala	Thr	Asp	Phe	Pro	Gly	Gln	Val	
				405					410						415	
caa	tgt	att	ttt	ctc	cg	aac	acc	agc	gcc	acc	gac	tca	ggg	gac	aaa	1296
Gln	Cys	Ile	Phe	Leu	Arg	Asn	Thr	Ser	Ala	Thr	Asp	Ser	Gly	Asp	Lys	
			420					425					430			
ttc	cca	tac	gat	acc	tct	ggg	ttt	aag	aag	ctc	aat	cag	tca	aac	tac	1344
Phe	Pro	Tyr	Asp	Thr	Ser	Gly	Phe	Lys	Lys	Leu	Asn	Gln	Ser	Asn	Tyr	
		435				440						445				
atg	ttc	ttc	ttg	cat	ccc	gat	gac	ttg	acg	aat	cta	gac	atc	gcg	aat	1392
Met	Phe	Phe	Leu	His	Pro	Asp	Asp	Leu	Thr	Asn	Leu	Asp	Ile	Ala	Asn	
		450				455					460					
ggc	cag	tgt	tac	aac	aca	tcc	att	ccg	cag	aat	tta	acg	ttt	agt	tat	1440
Gly	Gln	Cys	Tyr	Asn	Thr	Ser	Ile	Pro	Gln	Asn	Leu	Thr	Phe	Ser	Tyr	
465					470					475					480	
cag	ggt	ttg	cca	ttg	ggg	ctg	ggt	gat	gaa	ccc	aca	gct	gtc	aat	ggg	1488
Gln	Gly	Leu	Pro	Leu	Gly	Leu	Gly	Asp	Glu	Pro	Thr	Ala	Val	Asn	Gly	
				485					490						495	
tct	gcc	aat	cat	act	gct	gag	tcg	gcg	gcc	ggt	ttg	tcg	gta	aag	ggt	1536
Ser	Ala	Asn	His	Thr	Ala	Glu	Ser	Ala	Ala	Gly	Leu	Ser	Val	Lys	Gly	
			500					505					510			
ggg	act	gat	atg	cag	ggt	ttg	atc	ttt	ctg	ttc	acc	ctg	ggt	acg	gcc	1584
Gly	Thr	Asp	Met	Gln	Gly	Leu	Ile	Phe	Leu	Phe	Thr	Leu	Val	Thr	Ala	
		515					520					525				
tat	tcc	att	ttc	ttg	taa											1602
Tyr	Ser	Ile	Phe	Leu												
		530														

<210> 27

<211> 1392

<212> DNA

<213> Aspergillus niger

<220>

<221> CDS

<222> (1) .. (1392)

```

<400> 27
atg tcg aag tcg aat ggc gtc agc ctg gca ttc ccc gcc gaa gca gcg      48
Met Ser Lys Ser Asn Gly Val Ser Leu Ala Phe Pro Ala Glu Ala Ala
1      5      10      15
aca aag gaa tat gca gca tct cta gac tct tct gat aga ctt gct gca      96
Thr Lys Glu Tyr Ala Ala Ser Leu Asp Ser Ser Asp Arg Leu Ala Ala
      20      25      30
ttt cgg gag aag ttc atc gtc cca tcg aaa gca aat att gct tcc aaa      144
Phe Arg Glu Lys Phe Ile Val Pro Ser Lys Ala Asn Ile Ala Ser Lys
      35      40      45
aag cta gca aag cct ggc ctt tcg ccc gag tca tgc att tac ttc tgc      192
Lys Leu Ala Lys Pro Gly Leu Ser Pro Glu Ser Cys Ile Tyr Phe Cys
      50      55      60
ggg aac tca ctt ggc att caa ccc aag gcc aca gca aag tac tta gaa      240
Gly Asn Ser Leu Gly Ile Gln Pro Lys Ala Thr Ala Lys Tyr Leu Glu
      65      70      75      80
gca cag ttg gac act tgg tca tct att gga gtt ggc ggc cac ttc act      288
Ala Gln Leu Asp Thr Trp Ser Ser Ile Gly Val Gly Gly His Phe Thr
      85      90      95
gat ctt gag ggc tca ccc ttg aag caa tgg cag ctg ctc tca gag caa      336
Asp Leu Glu Gly Ser Pro Leu Lys Gln Trp Gln Leu Leu Ser Glu Gln
      100      105      110
gca gct gat tca atg agc aag att gtg gga gca aag cca gaa gaa gtc      384
Ala Ala Asp Ser Met Ser Lys Ile Val Gly Ala Lys Pro Glu Glu Val
      115      120      125
gca gcg atg gga aca ctt aca acg aac ctt cat ctt ctg ctg gct agt      432
Ala Ala Met Gly Thr Leu Thr Thr Asn Leu His Leu Leu Leu Ala Ser
      130      135      140
ttc tat aaa ccg act caa aca aag cac aag atc ttg atg gat tgg aag      480
Phe Tyr Lys Pro Thr Gln Thr Lys His Lys Ile Leu Met Asp Trp Lys
      145      150      155      160
gct ttc cca agt gat cac tac gct atc gaa tct cat att gcc tgg cat      528
Ala Phe Pro Ser Asp His Tyr Ala Ile Glu Ser His Ile Ala Trp His
      165      170      175
gac ctg gat cct aag gag tct atg gtg ctc atc ggt cca gat gaa ggc      576
Asp Leu Asp Pro Lys Glu Ser Met Val Leu Ile Gly Pro Asp Glu Gly
      180      185      190
gaa tac gag ata tcc acc cag aag att ttc tct tat atc gac aag cat      624
Glu Tyr Glu Ile Ser Thr Gln Lys Ile Phe Ser Tyr Ile Asp Lys His
      195      200      205
gct gat gag gct gct atg att ctt ctt cct ggt att caa tat tat act      672
Ala Asp Glu Ala Ala Met Ile Leu Leu Pro Gly Ile Gln Tyr Tyr Thr
      210      215      220
ggg caa ctg ttc gat atc cag aaa att aca aaa tat gcc cat tcg cgc      720
Gly Gln Leu Phe Asp Ile Gln Lys Ile Thr Lys Tyr Ala His Ser Arg
      225      230      235      240
aac atg gtt gtt ggc tgg gat ctg gcc cat gcg ttc gct aat gtt gag      768
Asn Met Val Val Gly Trp Asp Leu Ala His Ala Phe Ala Asn Val Glu
      245      250      255
cta aag ctg cat gat tgg aac gtt gat ttt gca gca tgg tgc aca tat      816
Leu Lys Leu His Asp Trp Asn Val Asp Phe Ala Ala Trp Cys Thr Tyr
      260      265      270
aag tac gga aac gcc ggt cct gga gca atg ggg ggg ctt ttc gtg cat      864
Lys Tyr Gly Asn Ala Gly Pro Gly Ala Met Gly Gly Leu Phe Val His
      275      280      285
gaa caa cac gga gag gtc gat tac agc gcg gga gaa gat gca ccc aaa      912

```

Glu	Gln	His	Gly	Glu	Val	Asp	Tyr	Ser	Ala	Gly	Glu	Asp	Ala	Pro	Lys	
290						295					300					
ttc	cgc	cat	cgt	ctg	acc	ggg	tgg	tat	ggg	ggc	gat	cga	tcg	gta	aga	960
Phe	Arg	His	Arg	Leu	Thr	Gly	Trp	Tyr	Gly	Gly	Asp	Arg	Ser	Val	Arg	
305				310					315						320	
ttc	aag	atg	gac	aac	aag	ttc	aaa	cct	att	cct	ggg	gca	gga	gga	ttt	1008
Phe	Lys	Met	Asp	Asn	Lys	Phe	Lys	Pro	Ile	Pro	Gly	Ala	Gly	Gly	Phe	
			325					330						335		
cag	ata	tca	aat	cct	tcg	gcc	atc	gac	ctt	gcg	tgt	ctt	tgt	gcc	gct	1056
Gln	Ile	Ser	Asn	Pro	Ser	Ala	Ile	Asp	Leu	Ala	Cys	Leu	Cys	Ala	Ala	
			340					345						350		
tta	tcg	gtg	ttt	gat	gag	acg	tca	atg	gca	gat	ctt	cgc	aga	aaa	tca	1104
Leu	Ser	Val	Phe	Asp	Glu	Thr	Ser	Met	Ala	Asp	Leu	Arg	Arg	Lys	Ser	
		355					360				365					
ttg	aag	cta	aca	gca	tac	ctt	gag	ttc	ctt	ttg	ctc	agg	gat	tat	gaa	1152
Leu	Lys	Leu	Thr	Ala	Tyr	Leu	Glu	Phe	Leu	Leu	Leu	Arg	Asp	Tyr	Glu	
	370					375					380					
gaa	gaa	tct	agg	cca	ttc	agc	att	atc	acg	ccc	aag	gac	ccc	gag	gcg	1200
Glu	Glu	Ser	Arg	Pro	Phe	Ser	Ile	Ile	Thr	Pro	Lys	Asp	Pro	Glu	Ala	
385					390					395					400	
aga	ggc	gca	caa	ctg	agt	cta	ttg	ctc	aag	cct	ggg	cta	tta	cag	aat	1248
Arg	Gly	Ala	Gln	Leu	Ser	Leu	Leu	Leu	Lys	Pro	Gly	Leu	Leu	Gln	Asn	
			405						410					415		
gtt	gca	cag	aag	ttg	caa	gaa	gca	gga	ata	gtc	tgc	gat	aaa	cgg	gag	1296
Val	Ala	Gln	Lys	Leu	Gln	Glu	Ala	Gly	Ile	Val	Cys	Asp	Lys	Arg	Glu	
			420					425					430			
ccg	ggg	gtt	gtg	cgt	gtc	gct	ccg	gtt	cct	ttg	tac	aac	agc	ttc	agc	1344
Pro	Gly	Val	Val	Arg	Val	Ala	Pro	Val	Pro	Leu	Tyr	Asn	Ser	Phe	Ser	
		435				440						445				
gag	gtt	tgg	aca	ttt	gtg	aag	atc	ttc	aag	gat	gca	ctg	cag	cag	tga	1392
Glu	Val	Trp	Thr	Phe	Val	Lys	Ile	Phe	Lys	Asp	Ala	Leu	Gln	Gln		
	450					455					460					

<210> 28

<211> 1143

<212> DNA

<213> *Aspergillus niger*

<220>

<221> CDS

<222> (1) .. (1143)

<400> 28

atg	gat	gtg	act	aat	ttg	gct	gtc	gat	tgt	ccc	gtt	ata	gag	acc	ggg	48
Met	Asp	Val	Thr	Asn	Leu	Ala	Val	Asp	Cys	Pro	Val	Ile	Glu	Thr	Gly	
1			5					10					15			
tcc	ctt	gct	ttg	tgg	atg	gct	tcc	tcc	gaa	gtg	gac	cag	aag	gcc	ttg	96
Ser	Leu	Ala	Leu	Trp	Met	Ala	Ser	Ser	Glu	Val	Asp	Gln	Lys	Ala	Leu	
			20					25				30				
ggc	aac	ttt	gcg	gcc	ata	acg	ggc	ctc	gag	aac	ccg	ttc	ctg	tcc	gcg	144
Gly	Asn	Phe	Ala	Ala	Ile	Thr	Gly	Leu	Glu	Asn	Pro	Phe	Leu	Ser	Ala	
		35					40					45				
atg	ctg	tac	aag	att	ctg	ggg	ttg	tct	gtt	atg	tta	tcg	aag	aaa	ctc	192
Met	Leu	Tyr	Lys	Ile	Leu	Gly	Leu	Ser	Val	Met	Leu	Ser	Lys	Lys	Leu	
	50					55					60					
ctc	cgt	gca	cga	cga	ttg	cgg	cgc	cta	gac	ccc	acc	cga	gag	act	aaa	240

Leu 65	Arg	Ala	Arg	Arg	Leu 70	Arg	Arg	Leu	Asp 75	Pro	Thr	Arg	Glu	Thr	Lys 80	
tcg	ctt	cat	ctt	tat	tac	cat	atc	atc	tgg	ctt	tcg	cgt	gag	ggg	ctt	288
Ser	Leu	His	Leu	Tyr	Tyr	His	Ile	Ile	Trp	Leu	Ser	Arg	Glu	Gly	Leu	
				85					90						95	
ctg	ata	tta	gag	gag	ttc	gtc	ctg	cca	ctg	gtt	gag	gga	ttt	atg	gaa	336
Leu	Ile	Leu	Glu	Glu	Phe	Val	Leu	Pro	Leu	Val	Glu	Gly	Phe	Met	Glu	
			100					105					110			
ctc	aaa	att	tta	gct	tac	aag	ctg	cgc	gcc	tcc	ttt	tac	cat	ata	ttc	384
Leu	Lys	Ile	Leu	Ala	Tyr	Lys	Leu	Arg	Ala	Ser	Phe	Tyr	His	Ile	Phe	
		115					120					125				
gtc	cta	ttt	cag	aac	cag	cct	gct	gtc	cac	tcg	ccg	ggg	atc	gtt	agt	432
Val	Leu	Phe	Gln	Asn	Gln	Pro	Ala	Val	His	Ser	Pro	Gly	Ile	Val	Ser	
	130					135					140					
ctg	ccg	agt	agt	aca	cct	tta	tcc	aac	ggc	gtg	aca	gaa	gcg	gag	tca	480
Leu	Pro	Ser	Ser	Thr	Pro	Leu	Ser	Asn	Gly	Val	Thr	Glu	Ala	Glu	Ser	
145					150				155						160	
cca	tcc	aag	gat	tcc	aat	ttg	aga	ttt	tca	ttt	cag	ctg	aag	ccc	gac	528
Pro	Ser	Lys	Asp	Ser	Asn	Leu	Arg	Phe	Ser	Phe	Gln	Leu	Lys	Pro	Asp	
				165					170					175		
acg	ata	acg	gtt	tcc	ggg	aag	ccc	agt	tcc	gcc	tcg	gat	agt	gcg	cca	576
Thr	Ile	Thr	Val	Ser	Gly	Lys	Pro	Ser	Ser	Ala	Ser	Asp	Ser	Ala	Pro	
			180					185					190			
cga	ggc	aag	gtt	acg	cag	gca	ccc	ccg	ggc	ttt	gcg	ccg	gtt	caa	ccg	624
Arg	Gly	Lys	Val	Thr	Gln	Ala	Pro	Pro	Gly	Phe	Ala	Pro	Val	Gln	Pro	
		195					200					205				
ccc	aag	tcg	aac	tcg	gcg	ttt	ctc	ctc	cct	gct	ctc	gat	tat	acc	cct	672
Pro	Lys	Ser	Asn	Ser	Ala	Phe	Leu	Leu	Pro	Ala	Leu	Asp	Tyr	Thr	Pro	
	210					215					220					
aca	gct	acc	gca	tgc	ttc	aat	cat	gcc	gcc	ctc	ttg	gca	gat	caa	ttc	720
Thr	Ala	Thr	Ala	Cys	Phe	Asn	His	Ala	Ala	Leu	Leu	Ala	Asp	Gln	Phe	
225					230					235					240	
ctc	ccg	ggg	tcg	cac	ccg	ctc	cgt	ctg	tcc	atc	aag	ctt	gaa	ttt	gcc	768
Leu	Pro	Gly	Ser	His	Pro	Leu	Arg	Leu	Ser	Ile	Lys	Leu	Glu	Phe	Ala	
				245					250					255		
gcc	tac	ctc	tac	gat	tgc	cta	cat	gac	gct	cat	gcc	tgt	cga	cgg	ttg	816
Ala	Tyr	Leu	Tyr	Asp	Cys	Leu	His	Asp	Ala	His	Ala	Cys	Arg	Arg	Leu	
			260					265					270			
gcc	aag	caa	gcc	att	gcc	gat	gtg	tac	aat	gcg	caa	gag	ggc	atg	gac	864
Ala	Lys	Gln	Ala	Ile	Ala	Asp	Val	Tyr	Asn	Ala	Gln	Glu	Gly	Met	Asp	
		275				280						285				
gac	gag	agc	ttc	gag	gat	gca	gct	gag	att	gtg	ggg	att	ttg	ggc	aag	912
Asp	Glu	Ser	Phe	Glu	Asp	Ala	Ala	Glu	Ile	Val	Gly	Ile	Leu	Gly	Lys	
	290					295				300						
atg	gtg	aag	cgg	gga	agg	aac	aca	agc	agt	gct	gga	aat	agc	acc	acg	960
Met	Val	Lys	Arg	Gly	Arg	Asn	Thr	Ser	Ser	Ala	Gly	Asn	Ser	Thr	Thr	
305					310					315					320	
gct	gtg	aaa	acg	ccc	cag	gag	gag	cga	agt	gag	ggc	agt	cga	aca	cca	1008
Ala	Val	Lys	Thr	Pro	Gln	Glu	Glu	Arg	Ser	Glu	Gly	Ser	Arg	Thr	Pro	
				325					330					335		
tcc	tct	cag	aca	acg	ttg	aaa	agg	ccg	gcg	cga	gtg	ccg	aaa	agc	acc	1056
Ser	Ser	Gln	Thr	Thr	Leu	Lys	Arg	Pro	Ala	Arg	Val	Pro	Lys	Ser	Thr	
			340					345					350			
tcc	tca	ccg	gcg	cgc	gcg	acc	aag	ccg	acg	act	agc	gag	ggg	atg	tcc	1104
Ser	Ser	Pro	Ala	Arg	Ala	Thr	Lys	Pro	Thr	Thr	Ser	Glu	Gly	Met	Ser	
		355					360					365				
ccc	gct	gtg	ccc	gac	ccc	acc	atg	atg	aat	ccc	ata	tga				1143

Pro Ala Val Pro Asp Pro Thr Met Met Asn Pro Ile
 370 375 380

<210> 29
 <211> 1779
 <212> DNA
 <213> *Aspergillus niger*

<220>
 <221> CDS
 <222> (1) .. (1779)

<400> 29
 atg aga aat ttg gag tgt gtc tac aca aaa tca cga agg tat ccc tcg 48
 Met Arg Asn Leu Glu Cys Val Tyr Thr Lys Ser Arg Arg Tyr Pro Ser
 1 5 10 15
 tct atc ttt ccg aca ctc agg agg tcg tca act agt ccg aac acg gcg 96
 Ser Ile Phe Pro Thr Leu Arg Arg Ser Ser Thr Ser Pro Asn Thr Ala
 20 25 30
 tcc tct cct gac act aat gcc tca tct tcg cca tgc tcg gac tct agt 144
 Ser Ser Pro Asp Thr Asn Ala Ser Ser Ser Pro Cys Ser Asp Ser Ser
 35 40 45
 ggg acc cat gct ggc att tcc att ccc cac tca gat agt ccc gcc agt 192
 Gly Thr His Ala Gly Ile Ser Ile Pro His Ser Asp Ser Pro Ala Ser
 50 55 60
 gtg gac gaa tgt ggg aac aac ggc ctt ttc gac ggg gaa tct cta act 240
 Val Asp Glu Cys Gly Asn Asn Gly Leu Phe Asp Gly Glu Ser Leu Thr
 65 70 75 80
 caa tta ttg gtc tct ggg agc aaa gcc ata tca ggc caa cca cat atc 288
 Gln Leu Leu Val Ser Gly Ser Lys Ala Ile Ser Gly Gln Pro His Ile
 85 90 95
 gcc att aag cca cat gtt ctg tct acc tat ccc ctt gac aac ctc ctc 336
 Ala Ile Lys Pro His Val Leu Ser Thr Tyr Pro Leu Asp Asn Leu Leu
 100 105 110
 tca gac att gag gga ctc ctc gag cag cca gtc tct cat ttg gct caa 384
 Ser Asp Ile Glu Gly Leu Leu Glu Gln Pro Val Ser His Leu Ala Gln
 115 120 125
 agc aag gaa gat cta aca aaa ccg gac cgg atg ttt cga act cgc ctt 432
 Ser Lys Glu Asp Leu Thr Lys Pro Asp Arg Met Phe Arg Thr Arg Leu
 130 135 140
 cca gcg gtc tca ggt gag gac aat gag tat ctg cta agg aaa gga gct 480
 Pro Ala Val Ser Gly Glu Asp Asn Glu Tyr Leu Leu Arg Lys Gly Ala
 145 150 155 160
 ctg ctg ctc ttg cca cca gcc ttg cgc aac gaa cta ttg gcg gca tat 528
 Leu Leu Leu Leu Pro Pro Ala Leu Arg Asn Glu Leu Leu Ala Ala Tyr
 165 170 175
 gtc aac tac gtc cat ccg cta tta ccg att ctt gat ctt gga gac ttt 576
 Val Asn Tyr Val His Pro Leu Leu Pro Ile Leu Asp Leu Gly Asp Phe
 180 185 190
 ctc tca caa atc att gtt gga gat gat gct caa ttc aag tca cca cta 624
 Leu Ser Gln Ile Ile Val Gly Asp Asp Ala Gln Phe Lys Ser Pro Leu
 195 200 205
 tta tac caa gct gtc atg ttc gcg ggc agt att ttt gct cag cga gac 672
 Leu Tyr Gln Ala Val Met Phe Ala Gly Ser Ile Phe Ala Gln Arg Asp
 210 215 220
 ggc aca gac gaa cct gaa cag cta acg agg gct tta ttt gaa cgt aca 720

Gly 225	Thr	Asp	Glu	Pro	Glu 230	Gln	Leu	Thr	Arg	Ala 235	Leu	Phe	Glu	Arg	Thr 240	
aag	gtt	cta	tac	gag	ttt	gaa	tcc	gaa	tca	tgt	gcc	tac	act	caa	att	768
Lys	Val	Leu	Tyr	Glu 245	Phe	Glu	Ser	Glu	Ser	Cys 250	Ala	Tyr	Thr	Gln 255	Ile	
caa	gct	ctc	tta	ttg	atg	act	ctc	tgg	cat	gga	gat	atg	tca	cgc	cac	816
Gln	Ala	Leu	Leu	Leu 260	Met	Thr	Leu	Trp	His	Gly 265	Asp	Met	Ser	Arg 270	His	
aaa	ggc	ccg	tct	tac	tgg	ctt	gac	gtc	gcc	ttc	tcc	act	gcg	gaa	aga	864
Lys	Gly	Pro	Ser	Tyr 275	Trp	Leu	Asp 280	Val	Ala	Phe	Ser	Thr 285	Ala	Glu	Arg	
att	ggc	cta	ctc	cat	gat	gaa	gac	tgg	cca	tgc	gat	agc	cgt	tcc	ttc	912
Ile	Gly	Leu	Leu	His	Asp	Asp	Trp 295	Pro	Cys	Asp 300	Ser	Arg	Ser	Phe		
cga	tcc	gct	atg	tgg	tgt	tgc	ctt	tac	gta	cgt	gat	cga	aca	atc	tcg	960
Arg	Ser	Ala	Met	Trp	Cys 310	Cys	Leu	Tyr	Val	Arg 315	Asp	Arg	Thr	Ile 320	Ser	
ctt	ggg	tct	cgt	cgg	cct	cca	cgg	atg	cca	gtc	aac	aaa	tat	ccg	gac	1008
Leu	Gly	Ser	Arg	Arg 325	Pro	Pro	Arg	Met	Pro	Val 330	Asn	Lys	Tyr	Pro 335	Asp	
tca	att	ctc	cat	ctt	atg	agc	gac	gca	cct	cgg	gaa	tat	gac	gaa	ata	1056
Ser	Ile	Leu	His	Leu 340	Met	Ser	Asp	Ala	Pro	Arg 345	Glu	Tyr	Asp	Glu 350	Ile	
att	ctt	gag	agt	cta	ggt	agc	gct	tct	ctg	atg	ctt	aca	agc	act	tac	1104
Ile	Leu	Glu	Ser	Leu 355	Gly	Ser	Ala 360	Ser	Leu	Met	Leu	Thr 365	Ser	Thr	Tyr	
caa	gaa	caa	tcg	acc	atg	ctc	ttc	aag	gaa	ctt	gtc	aaa	ctc	tcc	tac	1152
Gln	Glu	Gln	Ser	Thr 370	Met	Leu	Phe 375	Lys	Glu	Leu 380	Val	Lys	Leu	Ser	Tyr	
tgc	gtc	ggg	gca	atc	tta	gat	tat	cta	tac	gaa	ggg	gcc	tgg	gtc	cga	1200
Cys	Val	Gly	Ala	Ile 385	Leu	Asp	Tyr 390	Leu	Tyr	Glu 395	Gly	Ala	Trp	Val 400	Arg	
ata	cca	aca	cgt	cgc	agc	agc	ttt	tat	tca	ctc	gcc	cct	aaa	tgc	agc	1248
Ile	Pro	Thr	Arg	Arg 405	Ser	Ser	Phe 410	Tyr	Ser	Leu	Ala	Pro	Lys	Cys 415	Ser	
atc	tcg	cca	tcc	atc	aga	cca	agc	tgt	gag	aaa	cta	ctt	tac	tct	tgg	1296
Ile	Ser	Pro	Ser	Ile 420	Arg	Pro	Ser 425	Cys	Glu	Lys	Leu	Leu	Tyr	Ser 430	Trp	
att	caa	ttc	ctc	cct	ctc	aac	gca	gtc	tac	cat	ccg	cca	cag	ttg	ccc	1344
Ile	Gln	Phe	Leu	Pro 435	Leu	Asn	Ala 440	Val	Tyr	His	Pro	Pro	Gln	Leu 445	Pro	
gcc	gat	gcg	gaa	ggg	gag	tca	gcc	gaa	aca	gtc	ttc	ctt	gtc	cac	aaa	1392
Ala	Asp	Ala	Glu	Gly 450	Glu	Ser	Ala 455	Glu	Thr	Val	Phe	Leu	Val	His 460	Lys	
gcc	ttc	ctt	tac	ttg	ctg	tac	cta	gcc	tcc	atg	gca	gtc	ata	tac	cgc	1440
Ala	Phe	Leu	Tyr	Leu 465	Leu	Tyr	Leu 470	Ala	Ser	Met	Ala	Val	Ile	Tyr 480	Arg	
aca	ggc	act	cat	agc	caa	caa	acc	agc	acg	gca	act	ccg	aac	atg	gaa	1488
Thr	Gly	Thr	His	Ser 485	Gln	Gln	Thr	Ser	Thr	Ala 490	Thr	Pro	Asn	Met 495	Glu	
ggt	cta	cgg	gca	tcg	acg	tca	caa	ata	aaa	aag	gtc	ctt	gcg	gag	ctg	1536
Gly	Leu	Arg	Ala	Ser 500	Thr	Ser	Gln 505	Ile	Lys	Lys	Val	Leu	Ala	Glu 510	Leu	
cag	ggc	atg	gga	ctg	ctt	cag	ttc	ctt	cca	ggg	gca	tca	gtc	act	atc	1584
Gln	Gly	Met	Gly	Leu 515	Leu	Gln	Phe 520	Leu	Pro	Gly	Ala	Ser	Val	Thr 525	Ile	
ctg	atg	ttc	gct	gtg	gag	gct	agt	ctg	ttg	gat	ctc	cag	ggc	tcg	gat	1632

```

Leu Met Phe Ala Val Glu Ala Ser Leu Leu Asp Leu Gln Gly Ser Asp
530          535          540
acc atg gtc cga agg cag gct atg tca aat cta tat gcc tgt gag gag      1680
Thr Met Val Arg Arg Gln Ala Met Ser Asn Leu Tyr Ala Cys Glu Glu
545          550          555          560
gct gcg tta cac ttg atg cag gct tat cct acg gct gag atg gcg gta      1728
Ala Ala Leu His Leu Met Gln Ala Tyr Pro Thr Ala Glu Met Ala Val
          565          570          575
ttg aag act agg aca gcc cgt gca aac ctt ctt ggg ttg gtt gag aca      1776
Leu Lys Thr Arg Thr Ala Arg Ala Asn Leu Leu Gly Leu Val Glu Thr
          580          585          590
tga
1779

```

```

<210> 30
<211> 266
<212> PRT
<213> Aspergillus niger

```

```

<400> 30
Met Pro Ile Ser Ile Pro Ser Ala Ser Ser Val His Asp Leu Phe Ser
1      5      10      15
Leu Lys Gly Lys Val Val Val Ile Thr Gly Ala Ser Gly Pro Arg Gly
20      25      30
Met Gly Ile Glu Ala Ala Arg Gly Cys Ala Glu Met Gly Ala Asn Ile
35      40      45
Ala Leu Thr Tyr Ser Ser Arg Pro Gln Gly Gly Glu Lys Asn Ala Glu
50      55      60
Glu Leu Arg Asn Thr Tyr Gly Val Lys Ala Lys Ala Tyr Gln Cys Asn
65      70      75      80
Val Gly Asp Trp Asn Ser Val Lys Lys Leu Val Asp Asp Val Leu Ala
85      90      95
Glu Phe Gly Gln Ile Asp Ala Phe Ile Ala Asn Ala Gly Lys Thr Ala
100     105     110
Ser Ser Gly Ile Leu Asp Gly Ser Val Glu Asp Trp Glu Glu Val Ile
115     120     125
Gln Thr Asp Leu Thr Gly Thr Phe His Cys Ala Lys Ala Val Gly Pro
130     135     140
His Phe Lys Gln Arg Gly Thr Gly Ser Phe Ile Ile Thr Ser Ser Met
145     150     155     160
Ser Gly His Ile Ala Asn Phe Pro Gln Glu Gln Thr Ser Tyr Asn Val
165     170     175
Ala Lys Ala Gly Cys Ile His Met Ala Arg Ser Leu Ala Asn Glu Trp
180     185     190
Arg Asp Phe Ala Arg Val Asn Ser Ile Ser Pro Gly Tyr Ile Asp Thr
195     200     205
Gly Leu Ser Asp Phe Val Asp Lys Lys Thr Gln Asp Leu Trp Met Ser
210     215     220
Met Ile Pro Met Gly Arg Asn Gly Asp Ala Lys Glu Leu Lys Gly Ala
225     230     235     240
Tyr Val Tyr Leu Ala Ser Asp Ala Ser Thr Tyr Thr Thr Gly Ala Asp
245     250     255
Leu Val Ile Asp Gly Gly Tyr Thr Val Arg
260     265

```

```

<210> 31

```

<211> 533

<212> PRT

<213> *Aspergillus niger*

<400> 31

```

Met Ser Gly Val Ser Phe Ala Arg Gln Ala Leu Leu Gln Ala Ala Leu
1      5      10      15
Leu Ser Ile Val Ser Ala Ile Pro Ala Pro Thr Ala Phe Val Asn Asn
      20      25      30
Val Ala Pro Pro Glu Pro Ile Ile Thr Pro Ser Pro Val Gln His Arg
      35      40      45
Pro Ser Arg Val Ala Gly Arg Asn Ile Leu Ser Asp Val Asp Ser Asp
      50      55      60
Ile Asn Ser Ile Leu Ser Gly Leu Gly Ser Asp Leu Pro Ser Trp Val
65      70      75      80
Ala Ser Gly Val Pro Asn Tyr Phe Gln Gly Phe Pro Thr Gly Asp Ala
      85      90      95
Val Val Ser Ser Leu Gly Leu Asn Ser Ala Glu Leu Ala Ala Leu Pro
      100      105      110
Thr Asn Val Leu Asn Ile Asp Pro Tyr Ala Asn Trp Thr Ser Ser Gly
      115      120      125
Trp Asn Val Arg Phe His Gly Asn Val Tyr Lys Gln Pro Asn Thr Ser
      130      135      140
Ile Ser Asp Leu Asn Asp Leu Ala Asp Val Phe Leu Gly Asn Thr Ser
145      150      155      160
Ile Ser Asp Leu Ser Glu Ser Glu Gln Lys Gln Ala Arg Asn Leu Thr
      165      170      175
Ala Glu Ile Leu Val Val Gln Gln Ala His Val Ala Val Asn Thr Ile
      180      185      190
His Leu Glu Pro Ala Pro Ser Gln Gly Ser Ser Gly Gln Ser Gly Gly
      195      200      205
Gly Gly Ser Ser Asn Thr Thr Gly Gly Thr Gln Asp Leu Thr Leu Pro
      210      215      220
Tyr Asn Thr Thr Val Glu Gly Asp Phe Asp Thr Phe Val Pro Ile Ser
225      230      235      240
Ser Asn Gly Leu Thr Ala Gly Asn Glu Thr Ser Ala Ile Gln Arg Leu
      245      250      255
Asn Val His Val Glu Gly Ala Thr Ile Gly Asn Ser Thr Ala Tyr Leu
      260      265      270
Val Pro Pro Thr Gly Leu Thr Val Val Ser Asp Ile Asp Asp Ile Leu
      275      280      285
Arg Val Thr Lys Ile Tyr Glu Pro Glu Gln Gly Leu Leu Asn Ser Phe
      290      295      300
Ala Arg Pro Phe Arg Pro Trp Glu Asn Met Pro Asp Ile Tyr Arg Asn
305      310      315      320
Trp Ser Ile Ser Leu Pro Asn Leu His Phe His Tyr Leu Thr Thr Thr
      325      330      335
Pro Glu Gln Val Thr Arg Asn Tyr Met Gln Phe Ile Tyr Asp Asn Tyr
      340      345      350
Pro Gly Gly Ser Phe Asp Thr Arg Pro Leu Asn Phe Ser Asp Val Ser
      355      360      365
Ala Thr Leu Ser Ile Arg Lys Phe Leu Leu Gln Lys Val Phe Glu Thr
      370      375      380
Phe Pro Gln Arg Lys Phe Ile Leu Ile Ala Asp Thr Ser Asn Ser Asp
385      390      395      400
Val Met Arg Asp Tyr Pro Glu Met Ala Thr Asp Phe Pro Gly Gln Val
      405      410      415

```


Gln Cys Ile Phe Leu Arg Asn Thr Ser Ala Thr Asp Ser Gly Asp Lys
 420 425 430
 Phe Pro Tyr Asp Thr Ser Gly Phe Lys Lys Leu Asn Gln Ser Asn Tyr
 435 440 445
 Met Phe Phe Leu His Pro Asp Asp Leu Thr Asn Leu Asp Ile Ala Asn
 450 455 460
 Gly Gln Cys Tyr Asn Thr Ser Ile Pro Gln Asn Leu Thr Phe Ser Tyr
 465 470 475 480
 Gln Gly Leu Pro Leu Gly Leu Gly Asp Glu Pro Thr Ala Val Asn Gly
 485 490 495
 Ser Ala Asn His Thr Ala Glu Ser Ala Ala Gly Leu Ser Val Lys Gly
 500 505 510
 Gly Thr Asp Met Gln Gly Leu Ile Phe Leu Phe Thr Leu Val Thr Ala
 515 520 525
 Tyr Ser Ile Phe Leu
 530

<210> 32

<211> 463

<212> PRT

<213> Aspergillus niger

<400> 32

Met Ser Lys Ser Asn Gly Val Ser Leu Ala Phe Pro Ala Glu Ala Ala
 1 5 10 15
 Thr Lys Glu Tyr Ala Ala Ser Leu Asp Ser Ser Asp Arg Leu Ala Ala
 20 25 30
 Phe Arg Glu Lys Phe Ile Val Pro Ser Lys Ala Asn Ile Ala Ser Lys
 35 40 45
 Lys Leu Ala Lys Pro Gly Leu Ser Pro Glu Ser Cys Ile Tyr Phe Cys
 50 55 60
 Gly Asn Ser Leu Gly Ile Gln Pro Lys Ala Thr Ala Lys Tyr Leu Glu
 65 70 75 80
 Ala Gln Leu Asp Thr Trp Ser Ser Ile Gly Val Gly Gly His Phe Thr
 85 90 95
 Asp Leu Glu Gly Ser Pro Leu Lys Gln Trp Gln Leu Leu Ser Glu Gln
 100 105 110
 Ala Ala Asp Ser Met Ser Lys Ile Val Gly Ala Lys Pro Glu Glu Val
 115 120 125
 Ala Ala Met Gly Thr Leu Thr Thr Asn Leu His Leu Leu Leu Ala Ser
 130 135 140
 Phe Tyr Lys Pro Thr Gln Thr Lys His Lys Ile Leu Met Asp Trp Lys
 145 150 155 160
 Ala Phe Pro Ser Asp His Tyr Ala Ile Glu Ser His Ile Ala Trp His
 165 170 175
 Asp Leu Asp Pro Lys Glu Ser Met Val Leu Ile Gly Pro Asp Glu Gly
 180 185 190
 Glu Tyr Glu Ile Ser Thr Gln Lys Ile Phe Ser Tyr Ile Asp Lys His
 195 200 205
 Ala Asp Glu Ala Ala Met Ile Leu Leu Pro Gly Ile Gln Tyr Tyr Thr
 210 215 220
 Gly Gln Leu Phe Asp Ile Gln Lys Ile Thr Lys Tyr Ala His Ser Arg
 225 230 235 240
 Asn Met Val Val Gly Trp Asp Leu Ala His Ala Phe Ala Asn Val Glu
 245 250 255
 Leu Lys Leu His Asp Trp Asn Val Asp Phe Ala Ala Trp Cys Thr Tyr

<400>	33															
Met	Asp	Val	Thr	Asn	Leu	Ala	Val	Asp	Cys	Pro	Val	Ile	Glu	Thr	Gly	
1				5					10					15		
Ser	Leu	Ala	Leu	Trp	Met	Ala	Ser	Ser	Glu	Val	Asp	Gln	Lys	Ala	Leu	
			20					25					30			
Gly	Asn	Phe	Ala	Ala	Ile	Thr	Gly	Leu	Glu	Asn	Pro	Phe	Leu	Ser	Ala	
		35					40					45				
Met	Leu	Tyr	Lys	Ile	Leu	Gly	Leu	Ser	Val	Met	Leu	Ser	Lys	Lys	Leu	
	50					55					60					
Leu	Arg	Ala	Arg	Arg	Leu	Arg	Arg	Leu	Asp	Pro	Thr	Arg	Glu	Thr	Lys	
65				70						75					80	
Ser	Leu	His	Leu	Tyr	Tyr	His	Ile	Ile	Trp	Leu	Ser	Arg	Glu	Gly	Leu	
			85						90					95		
Leu	Ile	Leu	Glu	Glu	Phe	Val	Leu	Pro	Leu	Val	Glu	Gly	Phe	Met	Glu	
			100					105					110			
Leu	Lys	Ile	Leu	Ala	Tyr	Lys	Leu	Arg	Ala	Ser	Phe	Tyr	His	Ile	Phe	
		115					120					125				
Val	Leu	Phe	Gln	Asn	Gln	Pro	Ala	Val	His	Ser	Pro	Gly	Ile	Val	Ser	
	130					135					140					
Leu	Pro	Ser	Ser	Thr	Pro	Leu	Ser	Asn	Gly	Val	Thr	Glu	Ala	Glu	Ser	
145				150						155					160	
Pro	Ser	Lys	Asp	Ser	Asn	Leu	Arg	Phe	Ser	Phe	Gln	Leu	Lys	Pro	Asp	
				165					170					175		
Thr	Ile	Thr	Val	Ser	Gly	Lys	Pro	Ser	Ser	Ala	Ser	Asp	Ser	Ala	Pro	
			180					185					190			

```

Arg Gly Lys Val Thr Gln Ala Pro Pro Gly Phe Ala Pro Val Gln Pro
    195                200                205
Pro Lys Ser Asn Ser Ala Phe Leu Leu Pro Ala Leu Asp Tyr Thr Pro
    210                215                220
Thr Ala Thr Ala Cys Phe Asn His Ala Ala Leu Leu Ala Asp Gln Phe
    225                230                235                240
Leu Pro Gly Ser His Pro Leu Arg Leu Ser Ile Lys Leu Glu Phe Ala
    245                250                255
Ala Tyr Leu Tyr Asp Cys Leu His Asp Ala His Ala Cys Arg Arg Leu
    260                265                270
Ala Lys Gln Ala Ile Ala Asp Val Tyr Asn Ala Gln Glu Gly Met Asp
    275                280                285
Asp Glu Ser Phe Glu Asp Ala Ala Glu Ile Val Gly Ile Leu Gly Lys
    290                295                300
Met Val Lys Arg Gly Arg Asn Thr Ser Ser Ala Gly Asn Ser Thr Thr
    305                310                315                320
Ala Val Lys Thr Pro Gln Glu Glu Arg Ser Glu Gly Ser Arg Thr Pro
    325                330                335
Ser Ser Gln Thr Thr Leu Lys Arg Pro Ala Arg Val Pro Lys Ser Thr
    340                345                350
Ser Ser Pro Ala Arg Ala Thr Lys Pro Thr Thr Ser Glu Gly Met Ser
    355                360                365
Pro Ala Val Pro Asp Pro Thr Met Met Asn Pro Ile
    370                375                380

```

<210> 34

<211> 592

<212> PRT

<213> Aspergillus niger

<400> 34

```

Met Arg Asn Leu Glu Cys Val Tyr Thr Lys Ser Arg Arg Tyr Pro Ser
1      5      10      15
Ser Ile Phe Pro Thr Leu Arg Arg Ser Ser Thr Ser Pro Asn Thr Ala
    20      25      30
Ser Ser Pro Asp Thr Asn Ala Ser Ser Ser Pro Cys Ser Asp Ser Ser
    35      40      45
Gly Thr His Ala Gly Ile Ser Ile Pro His Ser Asp Ser Pro Ala Ser
    50      55      60
Val Asp Glu Cys Gly Asn Asn Gly Leu Phe Asp Gly Glu Ser Leu Thr
    65      70      75      80
Gln Leu Leu Val Ser Gly Ser Lys Ala Ile Ser Gly Gln Pro His Ile
    85      90      95
Ala Ile Lys Pro His Val Leu Ser Thr Tyr Pro Leu Asp Asn Leu Leu
    100     105     110
Ser Asp Ile Glu Gly Leu Leu Glu Gln Pro Val Ser His Leu Ala Gln
    115     120     125
Ser Lys Glu Asp Leu Thr Lys Pro Asp Arg Met Phe Arg Thr Arg Leu
    130     135     140
Pro Ala Val Ser Gly Glu Asp Asn Glu Tyr Leu Leu Arg Lys Gly Ala
    145     150     155     160
Leu Leu Leu Leu Pro Pro Ala Leu Arg Asn Glu Leu Leu Ala Ala Tyr
    165     170     175
Val Asn Tyr Val His Pro Leu Leu Pro Ile Leu Asp Leu Gly Asp Phe
    180     185     190
Leu Ser Gln Ile Ile Val Gly Asp Asp Ala Gln Phe Lys Ser Pro Leu

```

		195					200					205				
Leu	Tyr	Gln	Ala	Val	Met	Phe	Ala	Gly	Ser	Ile	Phe	Ala	Gln	Arg	Asp	
	210					215					220					
Gly	Thr	Asp	Glu	Pro	Glu	Gln	Leu	Thr	Arg	Ala	Leu	Phe	Glu	Arg	Thr	
225					230					235					240	
Lys	Val	Leu	Tyr	Glu	Phe	Glu	Ser	Glu	Ser	Cys	Ala	Tyr	Thr	Gln	Ile	
				245					250					255		
Gln	Ala	Leu	Leu	Leu	Met	Thr	Leu	Trp	His	Gly	Asp	Met	Ser	Arg	His	
			260					265					270			
Lys	Gly	Pro	Ser	Tyr	Trp	Leu	Asp	Val	Ala	Phe	Ser	Thr	Ala	Glu	Arg	
		275					280					285				
Ile	Gly	Leu	Leu	His	Asp	Glu	Asp	Trp	Pro	Cys	Asp	Ser	Arg	Ser	Phe	
	290					295					300					
Arg	Ser	Ala	Met	Trp	Cys	Cys	Leu	Tyr	Val	Arg	Asp	Arg	Thr	Ile	Ser	
305					310					315					320	
Leu	Gly	Ser	Arg	Arg	Pro	Pro	Arg	Met	Pro	Val	Asn	Lys	Tyr	Pro	Asp	
				325					330					335		
Ser	Ile	Leu	His	Leu	Met	Ser	Asp	Ala	Pro	Arg	Glu	Tyr	Asp	Glu	Ile	
			340					345					350			
Ile	Leu	Glu	Ser	Leu	Gly	Ser	Ala	Ser	Leu	Met	Leu	Thr	Ser	Thr	Tyr	
		355					360					365				
Gln	Glu	Gln	Ser	Thr	Met	Leu	Phe	Lys	Glu	Leu	Val	Lys	Leu	Ser	Tyr	
	370					375					380					
Cys	Val	Gly	Ala	Ile	Leu	Asp	Tyr	Leu	Tyr	Glu	Gly	Ala	Trp	Val	Arg	
385					390					395					400	
Ile	Pro	Thr	Arg	Arg	Ser	Ser	Phe	Tyr	Ser	Leu	Ala	Pro	Lys	Cys	Ser	
				405					410					415		
Ile	Ser	Pro	Ser	Ile	Arg	Pro	Ser	Cys	Glu	Lys	Leu	Leu	Tyr	Ser	Trp	
			420					425					430			
Ile	Gln	Phe	Leu	Pro	Leu	Asn	Ala	Val	Tyr	His	Pro	Pro	Gln	Leu	Pro	
		435					440					445				
Ala	Asp	Ala	Glu	Gly	Glu	Ser	Ala	Glu	Thr	Val	Phe	Leu	Val	His	Lys	
	450					455					460					
Ala	Phe	Leu	Tyr	Leu	Leu	Tyr	Leu	Ala	Ser	Met	Ala	Val	Ile	Tyr	Arg	
465					470					475					480	
Thr	Gly	Thr	His	Ser	Gln	Gln	Thr	Ser	Thr	Ala	Thr	Pro	Asn	Met	Glu	
				485					490					495		
Gly	Leu	Arg	Ala	Ser	Thr	Ser	Gln	Ile	Lys	Lys	Val	Leu	Ala	Glu	Leu	
			500					505					510			
Gln	Gly	Met	Gly	Leu	Leu	Gln	Phe	Leu	Pro	Gly	Ala	Ser	Val	Thr	Ile	
		515					520					525				
Leu	Met	Phe	Ala	Val	Glu	Ala	Ser	Leu	Leu	Asp	Leu	Gln	Gly	Ser	Asp	
	530					535					540					
Thr	Met	Val	Arg	Arg	Gln	Ala	Met	Ser	Asn	Leu	Tyr	Ala	Cys	Glu	Glu	
545					550				555							

```
<210> 35
<211> 1076
<212> DNA
<213> Artificial sequence
```

 $\langle 220 \rangle$

<221> PRIMER_BIND

<222> (1) .. (35)

<220>

<221> PROMOTER

<222> (13) .. (1062)

<220>

<221> PRIMER_BIND

<222> (1037) .. (1076)

<400> 35

ACTTCACTCG	AGATAATGTA	GGCGACAAAG	TAGCCGGGCA	TGCGACCGGA	GGA	CTCGGTC	60
AGAGGCACGA	ATATGAGAGG	CCAGAACGCC	GCACCCATGT	TCCAAGATGT	TGTGGCCCAG		120
TAAAGGTTAG	GAAATGGGGT	GTTTTGAACG	TGGAATCGTT	CTTCCATCCA	GTTGTTGCCG		180
GAGCCGTAGG	AACCGGCCGG	GATACCACTG	AGGATGGAAC	TATACCTATT	AGCACTGCCA		240
TCTTCGAGAA	CAAGGCTAAG	ACATACATGA	AACAAACGGT	GACCGTGAGA	AGCCATTTAT		300
AAGAGGTGCT	CCAATTGAAC	CTGAATATGT	TAAAACTAGT	ATACTAGCAA	CTGCATGCTT		360
GACGTACGGA	TTGTCCGGGT	CATCCTTGCC	GTTCCAGGTA	TGGATCTCAA	GGTCCTCCCT		420
GTCACGCAGA	TCATTGTGGC	TGACTACCCC	GACATGGGGG	TCCCGGCGAA	AGAACGACCG		480
CACCCTGCAA	CGTCACTTAG	TGGCTGTCTG	ACAGGGATTG	ATTGGCAATG	ATTGGCAGAC		540
TAACCGTCCC	AGAAACGAAT	CATCGAACTC	ATTGGTGTGC	ACTGTGCTAG	CGCGGCGCAG		600
CGACGGGCGT	GGCTCCTCGG	CAGGGGTGGA	CATGACTTTC	TCTGTGTAGA	TTCCTTGCAA		660
GGTATCTGCT	TGACAGATGC	AGAGATGATG	TTGAGTAAAA	TAGGGTGCAG	AGACCCTGCA		720
GAACCTGCAT	CGGCCGGCGG	CGACATGTCT	TCATGACCGA	CCCGGGCCGA	GCGGAATTAA		780
TGAAACTCGG	AACAGCCAGG	GGAATCCGAC	GCTTTGAGTG	GCCAATCCAT	GTCTATCCAC		840
ATCTGCATTG	AAGTGGATGA	AAGGTGGAGG	AGGGCTGGTT	GCCAAATTCT	GTTAGGGCTT		900
GGAATGTAAG	CAACCGGTTT	CGGTCATGAC	ATCATCACAG	CGATGACTTC	ATCACTCTCG		960
GACGAGCATA	TATAGTTGTG	CCTGTCGTCT	TATCTCAACA	AACATCAACA	ACAACAACAA		1020
TCATCTACTG	CATTTACCAA	ACTCATCTCT	ACCTCAATCA	ACTTAATTAA	TGAAGT		1076

<210> 36

<211> 1057

<212> DNA

<213> Artificial sequence

<220>

<221> PRIMER_BIND

<222> (1) .. (33)

<220>

<221> PROMOTER

<222> (13) .. (1043)

<220>

<221> PRIMER_BIND

<222> (1022) .. (1057)

<400> 36

ACTTCACTCG	AGTTCTCACA	AGCAAATCCG	AGAAACGCAA	TTGACCATAT	GTCCTTAGTA	60
TGAGGGCTAG	CAGGAATACT	TGTTTGTCGA	TGCATACCGA	CTAATAAAAA	TGGTTCTTCT	120
TTTTTCGCGCA	GGGATTTGGT	TTTTTCATTTT	TCATTTTTTCA	TATTTTATGT	TATTTTATTT	180
TATTTTTTTTG	TTTTATTTTA	TTTTATTTTT	TTTTTTTTTAA	TAAGACTGGG	TGTTTTCGTC	240
TCCCACCCCT	GCTGAGGAAC	TGTTTGATAT	ACAGAAGCCT	GAGATACTTG	AGATTATGTG	300
ACAAAGAAAT	TGTAAACCTG	AATGTGACTA	AGGCAGAAGA	GAGCAGGTGA	GGCCACACAA	360
TACTGTACAC	CCAATTCCGA	TCGTCATCGA	AATGAGCGTT	GTGGGTACTG	GAACGACAAA	420

CCATAGATCG	GATCTGCATC	CGAATAGGCT	TTACTTTGTA	CATTGGTAGG	AGTTTGAATG	480
TTTGAAGATC	CGGCAGATGA	TCACCTGCTG	CAGGGAATAA	TTTGAAACCT	GTCCGTCTGC	540
TTTGACCCAA	TTGACGAGGT	ACGTGATTTG	ATGCTACCCA	AGCAGTATGA	ATGCGACGGA	600
GGTGGCTATC	CTTACCTTGT	CATATGCAAC	TCATAACGGG	TAGTCCCCAT	GGAGCCACGG	660
CTATTTATTT	CTAAGGGCCA	ACAGTGAAAG	ATTGTGAGAT	CGTATACCTC	CTGGGTATGA	720
CGCAACCAGG	GTGAGCACTG	GCATTGCTCC	AATCCATATG	GTGGAGGTAT	GTCTCCGAGA	780
CCATGACGCA	TGTAGACCTT	GACCTGTCCC	TGATCCAGGG	ATCTGGTTTG	TCACTTGGCG	840
TATTTTAATT	AACTGCGGGT	CAGGGTCTTG	CCGTCCAGAC	CCAATGGATG	AATTGCCTGC	900
GTGGCCCTGC	AGTTCCAGCG	ATCGGCACGG	CACGGCTATG	ACGTCGGTGG	GTGTTCCGTT	960
TGATATTTAA	CCATAGAAGA	CAGCCTTGTT	ATGCTGGTTC	CTGCCGCAA	ATAAAGGTCT	1020
ATCCTTACCA	CGTTAAGTGT	CAAATTAATT	AAGAAGT			1057

<210> 37

<211> 1061

<212> DNA

<213> Artificial sequence

<220>

<221> PRIMER_BIND

<222> (1) .. (33)

<220>

<221> PROMOTER

<222> (13) .. (1047)

<220>

<221> PRIMER_BIND

<222> (1023) .. (1061)

<400> 37

ATCCTACTCG	AGTGGTGCCCT	GTGAACGAGG	TCACCACTGA	TAAATTGTCT	GGATCCGTAA	60
CGTCTGTATC	AGTGGATGGT	ACTGAAAAGT	TCCAAAGATT	GAGGGCAGCA	TATGTAGCAC	120
ATGGTGGTAG	CTCTAAGTGG	TTCGAAACTT	CGATGAATGG	GCGTGATATA	CAGGGGGGTA	180
GCACCTGTAT	AGTTGATGTC	AGCGAAGGCA	CGGGCCATGC	AGTGATAGAC	AAGAAAAGAG	240
GGATTGAATA	AAAGACTCGA	TTCTGAAGGG	GAGTAATACA	TCTTTCGGCT	TCTCACCACC	300
CCAGATGTAG	GCGTGTGTAA	GGTAAGCAAG	CACAACATAT	GCTCTCCGCC	ACTCAGGTTC	360
CCCCTTCAAC	TTCTTCGTAG	AGAGTATACC	CAGGTCATCG	ACAGACCTCC	TGATAGTACC	420
AGACTGAATC	AGGCCAGGCA	AGTCCTGAGC	GATTTCTCTC	CACGGGGAAT	AGTAAGGGTC	480
TTCCAATCTC	CGTACGGGAG	GAACCTCCGG	GAGGAACCCG	TTCTGGAGGG	AGACACCGTA	540
CGCAACAAGT	GAAATGTCTT	GTGTATAAAG	CATGATAATA	TATGCACTTC	AGTAGACAAC	600
TTGATATTCA	AGCCTCCTTG	AAATGATCCT	CATGCAGGTA	AGAGTAGATC	CAAGTGTTTA	660
TATTGTATT	CCAAGGAGGG	ACGGAAGAAA	ACCGGATCCA	GAGGGCTCGA	GATAGGGCCA	720
CCTGCATGAG	CAAAGACGGG	TCATTTCCCG	TTTCACAACG	ACCATTCTCA	GGCAGAAAGG	780
CAGTGGATTA	GCGGAAAAC	TAGCAACCAT	CCGGATTTCT	GGAGTGGATG	CTTTGTTATC	840
TCACTTCCGG	CAACACGGAT	GGTGAGGCGG	AAGCGGAACG	GTATCCGGAT	CCGGAGGGAG	900
GGACAAACCG	AAGTGCCTTG	TGTCACACGC	CGGTCCATAT	ATAAATGTTA	ATTGGTGACG	960
GTCAGCGGAT	TGTTTCTGGT	GTTTCATCCAT	TTTTTCTTGA	TATCCTGGAC	CCAGTTATTG	1020
CACTTTAATT	ACTATCCAAA	ATCCACATTA	ATTAATGAAG	T		1061

<210> 38

<211> 1097

<212> DNA

<213> Artificial sequence

<220>

<221> PRIMER_BIND

<222> (1) .. (34)

<220>

<221> PROMOTER

<222> (13) .. (1083)

<220>

<221> PRIMER_BIND

<222> (1060) .. (1097)

<400> 38

ACTTCACTCG	AGGTCCACAC	TCTAATTGGG	ATGAGCATTG	CCTTTCAGGA	ACTTGACTTT	60
TTGTGGTTGC	ATCTTTTGTT	GCCTTACCCA	CGCAGGCTTC	GTTTATGAGA	TAGCGATTTC	120
AGGCCTCAAA	TTGCTGTCAT	CACCTCCGGC	CCCAAACCTCG	TGGGCGGTGC	CCGCCCATCG	180
CTGGCAAGAA	GCACAGACTA	ACGCCTTACT	TCGCCTAGTT	GCACACGATG	CATCAAGACT	240
TCCTCCTATT	TCACAACGAC	CAGCTCCTTG	CCATTTGTTG	TTTCGAATGC	CCCCTTCCCC	300
GTGCCTTGCA	TTTCTAATTC	CCATCTATAC	CCTTTTGCGC	TCTCACGCCC	GCCTGACTAT	360
CAATTTTTCT	GCCAGCATAA	CCCTGCGTGC	AGAGCGGTTG	ACAGCAGGCT	TGCTAAGAGT	420
CTTCGTCTAG	AAGAAGTGTC	TGGGTGGATT	CCTTTTCGGA	TTAACTTTTG	CTATTTTAAA	480
AGCTGATTGC	ACTACGGAGT	ACACCCGCAA	TCTGATTCTC	ACTGATACTC	CAATTAAAGG	540
CCACTAATTG	CTGACATTTG	TAACATATTT	GCGTTTGCTT	TCCGGTTGAA	ATTGGCGGCT	600
TTGTGCTTTC	GATCGCTAGC	AGTCTCTGTG	CATGTACAAC	TGATCGGCTA	TTCAGTCCTA	660
TTCGGAAGCG	GCACTGAGAG	GCGGAGACAA	GCCACAGATA	CGACGAGTCT	GATTTTGAGA	720
CGGGCGCATC	CAACGAAAGC	CTGCGATTGC	AATTGAAGAT	CTCAATTGGG	GGACTGACAA	780
CCTAATACCT	AAACATATCG	AGGATTAGGC	GCCAACACGA	GCATGGGAAG	CTTTGGTGCA	840
TGCGACGGCT	GATCCATTTT	TAACAACCGT	TTTGCGCGCG	CGTCCGAGGT	AAGGCTCCCC	900
ATCCCTTGGC	GCTTTGCATG	TTTCACTCCG	ATGCATCTAC	TGTTTATATG	AGTGCGGTGG	960
GCAAGCTGTG	CCCACCGACT	GAGCATTCTC	AACCGGCCCA	ATGCTGTGGG	ATCGATCCTC	1020
CCCTCCAAAG	TTGTCGCTCG	GTTACACTTG	ATCGGAGCTT	GTTCTGTTAT	ACACCAAACC	1080
ATCTTAATTA	ATAGGAT					1097

<210> 39

<211> 1142

<212> DNA

<213> Artificial sequence

<220>

<221> PRIMER_BIND

<222> (1) .. (35)

<220>

<221> PROMOTER

<222> (13) .. (1128)

<220>

<221> PRIMER_BIND

<222> (1118) .. (1142)

<400> 39

ATCCTACTCG	AGGCCTTTGG	AGAATGTAAT	ATCCCTTGAT	CACCCCAGAC	TCACATCCAG	60
ATTCTGACTG	ACCTCGTGCA	TAGCCATATT	TAACCTGGTC	CCTGATGATC	AAAACAGTAC	120
ATGGAAGTGG	GGCCGAACAT	TTCCCTAGCC	AAAGGAGCAT	CAGGGATTCA	ACGTGATCAG	180
TACAGCCTTG	TAGCTCGCTG	AAACTGCCCT	GGAGCCAAGC	CAAAGCCTAG	ACAGCAAGGT	240
GCATGGGCGA	GCTTTGCGAG	GATTTAGTTG	TTATTTAGTA	TGCCAAGATT	AAGCCACGCA	300
GAATGGCTTA	TGGCCATAGT	AGCAACAACG	GCCACGGAAC	GTTCTTCCAT	GTGGTATGGG	360

GAAAGCCGCC	GGATATTTGG	GGAGGGCGTG	ATTGTTCCGG	TGGCTATCTT	TCCGGTGGAG	420
ATAACGGCGG	TGGACCATCA	AAACTCCTCA	GCCCGGATTG	CTCGGACTCC	GCATCGTCAA	480
GTATCGGTTTC	ATGACAGCC	ATTGACTAC	TTAGAAGACA	CAAATAGCTA	TCCAACTTAC	540
CCGGACCTTTG	ATAATAATT	TTCAGCAGAT	CTGTGCATCT	CAGAGACAAT	AATGCCACCG	600
TTCTGTGTGGA	AGGCACTGC	GCCGTGGTTT	AACAATAATA	AAACTCCAAA	TTGCCCAAAA	660
GCTAGAGCGAA	GGAAAGTCG	AGAAGCCCAG	TCTAGACTGG	CGGGCTGATC	GTCCGGTTCA	720
CAACCTGATTG	ACAAATCAT	CTGTGACCCT	CGATGTTCCG	CGGCGTTGAG	CGGGTGAGAA	780
TGTGGATCCTA	GTGGGGAAG	ACCCAATTGC	CGCCGGAAAA	GGTGAGCTTG	TGCTGCACAT	840
GTCCGGTGGCT	CAATTCAGG	GCCATGCCAG	GACATTGCTC	TTCAAAGGGA	TGCGATTGCT	900
ATTCCCTATTT	ACTTCTTTA	ACCCGCTGAC	GGGCCTGTCT	GGACCGAGAA	TTCTTGGTCA	960
AGTTTCAGTGG	AGCGATCCA	GGAAAGAACA	TCCGATATGG	ACAGTTGCCC	TCGATCTCAG	1020
GTCTGTAGTCG	TGTGACACA	GAGGCCTCTA	CTCGCTTGCC	TTGCATGTCT	GAGCTCAAAG	1080
GTGCGATGTGA	ATTAACACG	AGGAGCGACT	GTCTGTCTGCA	GATGCAACTT	AATTAATAGG	1140
AT						1142

<210> 40
 <211> 39
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 40	
ACTTCATTAA TTAAGTTGAT TGAGGTAGAG ATGAGTTTG	39

<210> 41
 <211> 35
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 41	
ACTTCACTCG AGATAATGTA GGCGACAAAG TAGCC	35

<210> 42
 <211> 36
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 42	
ACTTCATTAA TTAATTTGAC ACTTAACGTG GTAAGG	36

<210> 43
 <211> 33
 <212> DNA
 <213> Artificial sequence

<220>

<223> Primer

<400> 43

ACTTCACTCG AGTTCTCACA AGCAAATCCG AGA

33

<210> 44

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 44

ATCCTACTCG AGTGGTGCCT GTGAACGAGG TCA

33

<210> 45

<211> 39

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 45

ACTTCATTAA TTAATGTGGA TTTTGGATAG TAATTAAAG

39

<210> 46

<211> 38

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 46

ATCCTATTAA TTAAGATGGT TTGGTGTATA ACAGAACA

38

<210> 47

<211> 34

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 47

ACTTCACTCG AGGTCCACAC TCTAATTGGG ATGA

34

<210> 48

<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 48
ATCCTATTAA TTAAGTTGCA TCTGCGACAG ACAGT 35

<210> 49
<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 49
ACTTCACTCG AGGCCTTTGG AGAATGTAAT ATCCC 35

<210> 50
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 50
CAACACCTAC GGCGTCAAGG 20

<210> 51
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 51
TAACCCGGGG AGATGCTGTT 20

<210> 52
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 52
CGATCCGGAA ATCCTCCTG 20

<210> 53
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 53
GCCGCCGACT CAGCAGTAT 19

<210> 54
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 54
GTGCATGAAC AACACGGAGA 20

<210> 55
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 55
ACTCAGTTGT GCGCCTCTCG 20

<210> 56
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 56
GTCGAACTCG GCGTTTCTCC 20

<210> 57
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 57
CCGGTGAGGA GGTGCTTTTC 20

<210> 58
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 58
AGCATCTCGC CATCCATCAG 20

<210> 59
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 59
GCCTTCGGAC CATGGTATCC 20

<210> 60

<400> 60
000

<210> 61
<211> 1226
<212> DNA
<213> Artificial sequence

<220>
<221> gene
<222> (23) .. (739)
<223> GFP

<220>
<221> gene
<222> (740) .. (1212)
<223> BLE

<400> 61
GTCCTGTAA TTAACCTTCA CCATGGTGAG CAAGGGCGAG GAGCTGTTCA CCGGGGTGGT 60
GCCCATCCTG GTCGAGCTGG ACGGCGACGT AAACGGCCAC AAGTTCAGCG TGTCCGGCGA 120
GGGCGAGGGC GATGCCACCT ACGGCAAGCT GACCCTGAAG TTCATCTGCA CCACCGGCAA 180
GCTGCCCCGTG CCCTGGCCCA CCCTCGTGAC CACCCTGACC TACGGCGTGC AGTGCTTCAG 240
CCGCTACCCC GACCACATGA AGCAGCACGA CTTCTTCAAG TCCGCCATGC CCGAAGGCTA 300
CGTCCAGGAG CGCACCATCT TCTTCAAGGA CGACGGCAAC TACAAGACCC GCGCCGAGGT 360

GAAGTTCGAG	GGCGACACCC	TGGTGAACCG	CATCGAGCTG	AAGGGCATCG	ACTTCAAGGA	420
GGACGGCAAC	ATCCTGGGGC	ACAAGCTGGA	GTACAACTAC	AACAGCCACA	ACGTCTATAT	480
CATGGCCGAC	AAGCAGAAGA	ACGGCATCAA	GGTGAAC TTC	AAGATCCGCC	ACAACATCGA	540
GGACGGCAGC	GTGCAGCTCG	CCGACCACTA	CCAGCAGAAC	ACCCCCATCG	GCGACGGCCC	600
CGTGCTGCTG	CCCACAAACC	ACTACCTGAG	CACCCAGTCC	GCCCTGAGCA	AAGACCCCAA	660
CGAGAAGCGC	GATCACATGG	TCCTGCTGGA	GTTCGTGACC	GCCGCCGGGA	TCACTCTCGG	720
CATGGACGAG	CTGTACAAGG	CCAAGTTGAC	CAGTGCCGTT	CCGGTGCTCA	CCGCGCGCGA	780
CGTCGCCGGA	GCGGTCGAGT	TCTGGACCGA	CCGGCTCGGG	TTCTCCCGGG	ACTTCGTGGA	840
GGACGACTTC	GCCGGTGTGG	TCCGGGACGA	CGTGACCCTG	TTCATCAGCG	CGGTCCAGGA	900
CCAGGTGGTG	CCGGACAACA	CCCTGGCCTG	GGTGTGGGTG	CGCGGCCTGG	ACGAGCTGTA	960
CGCCGAGTGG	TCGGAGGTCG	TGTCCACGAA	CTTCCGGGAC	GCCTCCGGGC	CGGCCATGAC	1020
CGAGATCGGC	GAGCAGCCGT	GGGGGCGGGA	GTTCCGCCCTG	CGCGACCCGG	CCGGCAACTG	1080
CGTGCACTTC	GTGGCCGAGG	AGCAGGACTA	AAGGCGCGCC	TGAAGT		1226

<210> 62
 <211> 37
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 62
 GTCCTGTAA TTAACCTTCA CCATGGTGAG CAAGGGC 37

<210> 63
 <211> 36
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 63
 GCACTGGTCA ACTTGGCCTT GTACAGCTCG TCCATG 36

<210> 64
 <211> 36
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 64
 CATGGACGAG CTGTACAAGG CCAAGTTGAC CAGTGC 36

<210> 65
 <211> 35
 <212> DNA
 <213> Artificial sequence

<220>

<223> Primer

<400> 65

ACTTCAGGCG CGCCTTTAGT CCTGCTCCTC GGCCA

35

PCT

Print Out (Original in Electronic Form)

0-1	Form PCT/RO/134 (SAFE) Indications Relating to Deposited Microorganism(s) or Other Biological Material (PCT Rule 13bis)	
0-1-1	Prepared Using	
0-2	International Application No.	PCT/EP2006/060864
0-3	Applicant's or agent's file reference	24403WO

1	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
1-1	page	14
1-2	line	10
1-3	Identification of deposit	
1-3-1	Name of depositary institution	CBS Centraalbureau voor Schimmelcultures
1-3-2	Address of depositary institution	Uppsalalaan 8, NL-3584 CT Utrecht, The Netherlands / P.O. Box 85167, NL-3508 AD Utrecht, The Netherlands
1-3-3	Date of deposit	10 August 1998 (10.08.1998)
1-3-4	Accession Number	CBS 513.88
1-4	Additional Indications	Aspergillus Niger DS2975
1-5	Designated States for Which Indications are Made	all designations
1-6	Separate Furnishing of Indications	Form 134 added as separate indication concerning microorganism to be included in the specification
	These indications will be submitted to the International Bureau later	

FOR RECEIVING OFFICE USE ONLY

0-4	This form was received with the international application:	
	(yes or no)	
0-4-1	Authorized officer	

FOR INTERNATIONAL BUREAU USE ONLY

0-5	This form was received by the international Bureau on:	
0-5-1	Authorized officer	